

201359

STIC-Biotech/ChemLib

From: Ramirez, Delia
Sent: Tuesday, September 12, 2006 12:19 PM
To: STIC-Biotech/ChemLib
Subject: 10/806370

CRFE

Hi,

I would like to request the following searches:

1. SEQ ID NO: 11 and 12 in the protein databases (commercial and interference).
2. SEQ ID NO:11 in the nucleic acid databases (commercial and interference).

Please provide a printout of the results.

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

78701

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:43:56 ; Search time 9692 Seconds
(without alignments)
11533.233 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 1748
Sequence: 1 ctcagagaccatgagagaggg.....gacctgctgacccagagaa 1748

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_env: *
2: gb_pat: *
3: gb_ph: *
4: gb_pl: *
5: gb_pr: *
6: gb_ro: *
7: gb_sts: *
8: gb_sy: *
9: gb_un: *
10: gb_vi: *
11: gb_ov: *
12: gb_hcg: *
13: gb_in: *
14: gb_om: *
15: gb_da: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1748	100.0	1748	2	ARS34544
2	1748	100.0	1748	2	AX369323
3	1687.4	96.5	1689	2	CS075405
4	1687.4	96.5	1689	2	DD194635
5	1687.4	96.5	1689	2	AX149579
6	1670.4	95.6	2380	2	CS081489
7	1670.4	95.6	2393	2	DD200135
8	1670.4	95.6	2393	2	AB048796
9	1670.2	95.5	2237	5	AY190317
10	1663.4	95.2	2067	2	CS075313
11	1663.4	95.2	2067	2	DD194636
12	1663.4	95.2	2067	2	AX149581
13	1523.4	87.2	3222	5	BC018715
14	1479	84.6	3324	5	AB048797
15	1429	81.8	1614	2	AX360089
16	1302.8	74.5	1314	2	AX190543
17	1302.8	74.5	1314	2	AX190546
18	1277.8	73.1	1574	2	AX704683

19	1277	73.1	1375	2	CQ729940	CQ729940 Sequence
20	1258.6	72.0	1434	2	ARS83622	ARS83622 Sequence
21	1252.8	71.7	2192	2	BD156842	BD156842 Primer fo
22	1252.8	71.7	2192	2	AX877776	AX877776 Sequence
23	1252.8	71.7	2192	5	AK027798	AK027798 Homo sapi
24	1168	66.8	3320	6	BC085323	BC085323 Mus muscu
25	1111.2	63.6	2992	6	BC042878	BC042878 Mus muscu
26	1068.4	61.1	1078	2	AX190545	AX190545 Sequence
27	1062.6	60.8	1341	2	ARS83623	ARS83623 Sequence
28	1053	60.2	1230	2	ARS34553	ARS34553 Sequence
29	1053	60.2	1230	2	AX369347	AX369347 Sequence
30	782.4	44.8	1222	2	BD276278	BD276278 Serine pr
31	709.4	40.6	834	2	ARS34546	ARS34546 Sequence
32	709.4	40.6	834	2	AX369340	AX369340 Sequence
33	535.6	30.6	852	2	BD149569	BD149569 Primer fo
34	535.6	30.6	852	2	AX869507	AX869507 Sequence
35	446	25.5	2125	6	BC010843	BC010843 Mus muscu
36	440.8	25.2	678	2	ARS34548	ARS34548 Sequence
37	440.8	25.2	678	2	AX369342	AX369342 Sequence
38	411.6	23.5	614	2	ARS34551	ARS34551 Sequence
39	411.6	23.5	614	2	AX369345	AX369345 Sequence
40	394.8	22.6	160222	5	AP002962	AP002962 Homo sapi
41	394.8	22.6	184926	5	AP000757	AP000757 Homo sapi
42	348	19.9	402	2	ARS34534	ARS34534 Sequence
43	348	19.9	402	2	AX369313	AX369313 Sequence
44	335.8	19.2	621	2	ARS34547	ARS34547 Sequence
45	335.8	19.2	621	2	AX369341	AX369341 Sequence

ALIGNMENTS

RESULT 1
ARS34544
LOCUS ARS34544
DEFINITION Sequence 11 from patent US 6734006.
ACCESSION ARS34544
VERSION ARS34544.1 GI:53924854
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1748)
AUTHORS Xiao, Y. and Gedrich, R.W.
TITLE Regulation of human transmembrane serine protease
JOURNAL Patent: US 6734006-A 11 11-MAY-2004;
Bayer Aktiengesellschaft; Leverkusen;
WOX;

FEATURES
source
Location/Qualifiers
1..1748
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1748; DB 2; Length 1748;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCAGAGACCATGAGAGGAGGACGACCGGAGATGATCTTCAGCAAGAACCTTCAGC 60
1 CTTCAGAGACCATGAGAGGAGGACGACCGGAGATGATCTTCAGCAAGAACCTTCAGC 60
DB 1 CTTCAGAGACCATGAGAGGAGGACGACCGGAGATGATCTTCAGCAAGAACCTTCAGC 60
QY 61 TGGAGCATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 120
61 TGGAGCATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 120
DB 61 TGGAGCATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 120
61 TGGAGCATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 120
QY 121 CCAGGATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 180
121 CCAGGATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 180
DB 121 CCAGGATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 180
121 CCAGGATCTTCAGCCGAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGC 180
QY 181 CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGGCGGGGATCTTCAGC 240
181 CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGGCGGGGATCTTCAGC 240
DB 181 CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGGCGGGGATCTTCAGC 240
181 CCAGGATCTTCAGCTGGTACCTTCAGGCGGGGATCTTCAGGCGGGGATCTTCAGC 240

QY 241 CCAGGCATCTCAGCCCGGAGCATCTCCGCTCTGGCATCACTTTCCAGGCTTCATCCGG 300
 DB 241 CCAAGGCATCTCAGCCCGGAGCATCTCCGCTCTGGCATCACTTTCCAGGCTTCATCCGG 300
 QY 301 CAGGTCATCTCCGCAAGTCAGGCTCGGTGAACAATCTCCCAACAGAGTGTACTTGT 360
 DB 301 CAGGTCATCTCCGCAAGTCAGGCTCGGTGAACAATCTCCCAACAGAGTGTACTTGT 360
 QY 361 TAGAGCAACATCAGTGGGGGCTGTATCCCATCGATCATCTCTCGCAGGTGAGACCAAGC 420
 DB 361 TAGAGCAACATCAGTGGGGGCTGTATCCCATCGATCATCTCTCGCAGGTGAGACCAAGC 420
 QY 421 AACGAGGCGCACAGGAGGAGCCGAGGTACAGAGCTGGCCAGATTCACTTGGCGAGAGG 480
 DB 421 AACGAGGCGCACAGGAGGAGCCGAGGTACAGAGCTGGCCAGATTCACTTGGCGAGAGG 480
 QY 481 CCAAGAGAGCTACCGCTCATCGGGTGGTGTCTCTCTCATTTGGCCTGTGTGTCTGCT 540
 DB 481 CCAAGAGAGCTACCGCTCATCGGGTGGTGTCTCTCTCATTTGGCCTGTGTGTCTGCT 540
 QY 541 CATCATCTCTTCCAGTCTTGAGAGGCGCACAGAGGATCAGGTACAGAGGAGAGGAG 600
 DB 541 CATCATCTCTTCCAGTCTTGAGAGGCGCACAGAGGATCAGGTACAGAGGAGAGGAG 600
 QY 601 GAGGTGTCCAGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 DB 601 GAGGTGTCCAGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 QY 661 CGAGCTGGGCTGGCGTGAAGTTTGAATGGGAGCAAGTCTTGTCTTAAATCTACTCTGG 720
 DB 661 CGAGCTGGGCTGGCGTGAAGTTTGAATGGGAGCAAGTCTTGTCTTAAATCTACTCTGG 720
 QY 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATGACTCTTCTCAGAGAGAG 780
 DB 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATGACTCTTCTCAGAGAGAG 780
 QY 781 CTGCGCAGAGCTGGGTTTGAAGTGTCTCACCGGACCAAGGTTGGCCACAGAGATT 840
 DB 781 CTGCGCAGAGCTGGGTTTGAAGTGTCTCACCGGACCAAGGTTGGCCACAGAGATT 840
 QY 841 TGCCCAACAGCTTCTCAATCTTGAATACAACTCCACATCCAGAGAAAGCTTCCACAG 900
 DB 841 TGCCCAACAGCTTCTCAATCTTGAATACAACTCCACATCCAGAGAAAGCTTCCACAG 900
 QY 901 TGAATGCCCTTCCAGAGGTAATCTCCCTCAGATGCCACTGGGAGCTGAGAGGAGCAT 960
 DB 901 TGAATGCCCTTCCAGAGGTAATCTCCCTCAGATGCCACTGGGAGCTGAGAGGAGCAT 960
 QY 961 GACCGGGGAGATCGTGGAGGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGA 1020
 DB 961 GACCGGGGAGATCGTGGAGGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGA 1020
 QY 1021 TCTGCACTTGGGCAACCCCACTCTGTGAGGAGCAGCTCAATGACCGCCAGTGGGTGCT 1080
 DB 1021 TCTGCACTTGGGCAACCCCACTCTGTGAGGAGCAGCTCAATGACCGCCAGTGGGTGCT 1080
 QY 1081 CACTGCGGCCACTGCTTCTGTGAGCCGGGAGAGAGGTCTTGGAGGGCTGGAAGGTGA 1140
 DB 1081 CACTGCGGCCACTGCTTCTGTGAGCCGGGAGAGAGGTCTTGGAGGGCTGGAAGGTGA 1140
 QY 1141 CGCGGGGCGACAGCACTGTGACAGATTGCTGTGAGGAGAGCTTCAATGCGAGATCAT 1200
 DB 1141 CGCGGGGCGACAGCACTGTGACAGATTGCTGTGAGGAGAGCTTCAATGCGAGATCAT 1200
 QY 1201 CAAAGAGCAATTAACCGATGAGAGAGAGATTAAGACATGGCCCTCATGCGGCTGTCAA 1260
 DB 1201 CAAAGAGCAATTAACCGATGAGAGAGAGATTAAGACATGGCCCTCATGCGGCTGTCAA 1260
 QY 1261 GCGCCCTGACCTGTGCTGATCATCACTGCTTGTGCTTCCCATGAGTGAAGAGCTT 1320
 DB 1261 GCGCCCTGACCTGTGCTGATCATCACTGCTTGTGCTTCCCATGAGTGAAGAGCTT 1320

QY 1321 TAGCCTCAATGAGACTGTGTGATCAGAGGCTTTGGCAAGACCAAGGAGACAGATGACAA 1380
 DB 1321 TAGCCTCAATGAGACTGTGTGATCAGAGGCTTTGGCAAGACCAAGGAGACAGATGACAA 1380
 QY 1381 GACATCCCCCTCTCCGGGAGGTGACAGTCAATCTTATGACTTCAAGAAATGCAATGA 1440
 DB 1381 GACATCCCCCTCTCCGGGAGGTGACAGTCAATCTTATGACTTCAAGAAATGCAATGA 1440
 QY 1441 CTACTTGCTATGACAGATTACCTTACCCCAAGGATGATGTGTGCTGGGACCTTGTGG 1500
 DB 1441 CTACTTGCTATGACAGATTACCTTACCCCAAGGATGATGTGTGCTGGGACCTTGTGG 1500
 QY 1501 GGGGAGAGACTCTGCGAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAGAAACCG 1560
 DB 1501 GGGGAGAGACTCTGCGAGGAGAGACAGCGGGGGGCTCTTGTCTGTGAGAGAGAAACCG 1560
 QY 1561 CTGTTACTTGGCAGAGTTCACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG 1620
 DB 1561 CTGTTACTTGGCAGAGTTCACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG 1620
 QY 1621 TGTGTACACCAAGTGCAGAAAGTTCTTCCCTGGATTACAGCAAGATGAGAGAGAGT 1680
 DB 1621 TGTGTACACCAAGTGCAGAAAGTTCTTCCCTGGATTACAGCAAGATGAGAGAGAGT 1680
 QY 1681 GCGATTCAAGAAATCTTAACCAAGCTGGCTGTGCTCTGCAACAGACCGGCTGTGTGAC 1740
 DB 1681 GCGATTCAAGAAATCTTAACCAAGCTGGCTGTGCTCTGCAACAGACCGGCTGTGTGAC 1740
 QY 1741 TCGAGAAA 1748
 DB 1741 TCGAGAAA 1748

RESULT 2
 AX369323
 LOCUS AX369323 1748 bp DNA linear PAT 16-FEB-2002
 DEFINITION Sequence 11 from Patent WO0196538.
 ACCESSION AX369323
 VERSION AX369323.1 GI:18857263
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Xlao, Y. and Gedrich, R.
 TITLE Regulation of human transmembrane serine protease
 JOURNAL Patent: WO 0196538-A 11 20-DEC-2001;
 FEATURES
 location/Qualifiers
 1..1748
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1748; DB 2; Length 1748;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCAGAGACCATGAGAGGAGACAGCCAGGAAATGCAATCTTCAGAGAGAAACACCTTGAGC 60
 DB 1 CTCAGAGACCATGAGAGGAGACAGCCAGGAAATGCAATCTTCAGAGAGAAACACCTTGAGC 60
 QY 61 TGAAGCATCTCCAGCCCAAGGATCTCCAGTGGGACACTTCCAGGCGGGGACATCTCCAGC 120
 DB 61 TGAAGCATCTCCAGCCCAAGGATCTCCAGTGGGACACTTCCAGGCGGGGACATCTCCAGC 120
 QY 121 CCAGGCATCTCCAGCCCAAGGATCTCCAGTGGGACACTTCCAGGCGGGGACATCTCCAGC 180
 DB 121 CCAGGCATCTCCAGCCCAAGGATCTCCAGTGGGACACTTCCAGGCGGGGACATCTCCAGC 180

QY 181 CCAGGATCTTCAGCTGGTACACCTCCAGGCGGGGATCTCAAGCGGGGATCTCCAGC 240
 DB 181 CCAGGATCTTCAGCTGGTACACCTCCAGGCGGGGATCTTCAGGCGGGGATCTCCAGC 240
 QY 241 CCAGGATCTTCAGGCGGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCATCCG 300
 DB 241 CCAGGATCTTCAGGCGGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCATCCG 300
 QY 301 CAGGTATCATCCGCGGATCTAGCTGGTGAACAACCTCCCGCAACGAGGTATCTTTGT 360
 DB 301 CAGGTATCATCCGCGGATCTAGCTGGTGAACAACCTCCCGCAACGAGGTATCTTTGT 360
 QY 361 TAGAGCAACACGATGGGGGCTGTACCATCCATCTCTCTCCAGGATCAGACGAGC 420
 DB 361 TAGAGCAACACGATGGGGGCTGTACCATCCATCTCTCTCCAGGATCAGACGAGC 420
 QY 421 AACGAGGCGCCACGAGGAGAGCCGAGGTACGAGCTTCCCAAGTTCACTTGGCGGAGG 480
 DB 421 AACGAGGCGCCACGAGGAGAGCCGAGGTACGAGCTTCCCAAGTTCACTTGGCGGAGG 480
 QY 481 CCAGAACAGCTACCGCTCATTCGGGTGCTGCTCTCTCATTTGCGCTGGTGTGGCT 540
 DB 481 CCAGAACAGCTACCGCTCATTCGGGTGCTGCTCTCTCATTTGCGCTGGTGTGGCT 540
 QY 541 CATCATCTCTTCAGATTCTGGAGGCGACACAGGATCAGATCAAGAGAGAGAGG 600
 DB 541 CATCATCTCTTCAGATTCTGGAGGCGACACAGGATCAGATCAAGAGAGAGAGG 600
 QY 601 GAGCTGTCCAGAGACGCTGTTGCTGCTGAGCGGGGTGGTGAATGCAAGCTGAAGAGTGA 660
 DB 601 GAGCTGTCCAGAGACGCTGTTGCTGCTGAGCGGGGTGGTGAATGCAAGCTGAAGAGTGA 660
 QY 661 CGAGCTGGGCTGGCTGAGGTTGACTGGAGACAACTCTGTTAAATCTTACTCTGGGCT 720
 DB 661 CGAGCTGGGCTGGCTGAGGTTGACTGGAGACAACTCTGTTAAATCTTACTCTGGGCT 720
 QY 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGCAACTGGAAATGACTCTACTCAGAGAGAGC 780
 DB 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGCAACTGGAAATGACTCTACTCAGAGAGAGC 780
 QY 781 CTGGCAGCAGCTGGGTTTCCAGAGTGTCTCAACGAGCAACCGAGGTTGGCCACAGGATTT 840
 DB 781 CTGGCAGCAGCTGGGTTTCCAGAGTGTCTCAACGAGCAACCGAGGTTGGCCACAGGATTT 840
 QY 841 TGCCCAACAGCTTCTCAATCTTGAGATACATCTCAACATCCAGGAAAGCTTCCACAGGTC 900
 DB 841 TGCCCAACAGCTTCTCAATCTTGAGATACATCTCAACATCCAGGAAAGCTTCCACAGGTC 900
 QY 901 TGAATGCTTCCAGAGGTTATCTCCCTCAGGTTCCCATGCGGACTGAGGGCCAT 960
 DB 901 TGAATGCTTCCAGAGGTTATCTCCCTCAGGTTCCCATGCGGACTGAGGGCCAT 960
 QY 961 GACCGGCGGATCTGGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTAG 1020
 DB 961 GACCGGCGGATCTGGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTAG 1020
 QY 1021 TCTGCACTTGGGCAACCACTCTGTGAGGCACTCTCATTTGACCGCCAGTGGGTCT 1080
 DB 1021 TCTGCACTTGGGCAACCACTCTGTGAGGCACTCTCATTTGACCGCCAGTGGGTCT 1080
 QY 1081 CACTGCGCGCCACAGCTTCTGTGACCCGGAGAGAGGTCTTGAGGGGCTGGAGGTGA 1140
 DB 1081 CACTGCGCGCCACAGCTTCTGTGACCCGGAGAGAGGTCTTGAGGGGCTGGAGGTGA 1140
 QY 1141 CGCGGGGACAGCAACCTGCAACAGTTGCTGAGGCAAGCTTCAATGCGAGATCATCAT 1200
 DB 1141 CGCGGGGACAGCAACCTGCAACAGTTGCTGAGGCAAGCTTCAATGCGAGATCATCAT 1200
 QY 1201 CAAACGCAATTAACCAAGTGAAGAGAGATTAAGACATGCGCTCATGCGGCTGTCCAA 1260
 DB 1201 CAAACGCAATTAACCAAGTGAAGAGAGATTAAGACATGCGCTCATGCGGCTGTCCAA 1260
 QY 1261 GCCCTGACCTGTGCTGATCAATCAACCTGTGCTTCCCATGATGAGACAGACTT 1320

DB 1261 GCCCTGACCTGTGCTGATCAATCAACCTGTGCTTCCCATGATGAGACAGACTT 1320
 QY 1321 TAGCTTCAATGAACACCTGCTGATCAACAGGCTTTGGCAACAGAGAGAGATGACAA 1380
 DB 1321 TAGCTTCAATGAACACCTGCTGATCAACAGGCTTTGGCAACAGAGAGAGATGACAA 1380
 QY 1381 GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGA 1440
 DB 1381 GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGA 1440
 QY 1441 CTACTGTTATGACAGTTACTTACCCCAAGATGATGTGCTGGGACCTTCTGTG 1500
 DB 1441 CTACTGTTATGACAGTTACTTACCCCAAGATGATGTGCTGGGACCTTCTGTG 1500
 QY 1501 GGGCAGAGATCTCTCCAGAGGAGACACAGCGGGGCTCTTGTCTGTGAGAGAGAACACCG 1560
 DB 1501 GGGCAGAGATCTCTCCAGAGGAGACACAGCGGGGCTCTTGTCTGTGAGAGAGAACACCG 1560
 QY 1561 CTGGTACCTGGCAGAGTGTACACAGCTGGGCAACAGCTGTGGCCAGAGAACAACTGG 1620
 DB 1561 CTGGTACCTGGCAGAGTGTACACAGCTGGGCAACAGCTGTGGCCAGAGAACAACTGG 1620
 QY 1621 TGTGTACACCAAGTGAACAGAGTTCTTCCCTGATTTAAGCAAGATGAGAGCGAGT 1680
 DB 1621 TGTGTACACCAAGTGAACAGAGTTCTTCCCTGATTTAAGCAAGATGAGAGCGAGT 1680
 QY 1681 GCGATTGAGAAATCTTAACACAGCTGGCTGCTGTGACACACACCGCTGTGTGAC 1740
 DB 1681 GCGATTGAGAAATCTTAACACAGCTGGCTGCTGTGACACACACCGCTGTGTGAC 1740
 QY 1741 TCGAGAAA 1748
 DB 1741 TCGAGAAA 1748
 RESULT 3
 CS075405 1689 bp DNA linear PAT 05-MAY-2005
 LOCUS CS075405
 DEFINITION Sequence 93 from Patent WO2005019270.
 ACCESSION CS075405
 VERSION CS075405.1 GI:63091711
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE
 1. Madison, E.L. and Nixon, A.
 Endothelinase-2 ligands
 Patent: WO 2005019270-A 93 03-MAR-2005;
 JOURNAL Dyax Corp. (US); Dendreon Corporation (US)
 FEATURES
 source
 1. 1689
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 96.5%; Score 1687.4; DB 2; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 11 ATGAGAGGAGACAGCCAGGAGATGATCTTCAGCAAGAACACCTTCACTGAGCATCT 70
 DB 1 ATGAGAGGAGACAGCCAGGAGATGATCTTCAGCAAGAACACCTTCACTGAGCATCT 60
 QY 71 CCAGCCAGGATCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 130
 DB 61 CCAGCCAGGATCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 120
 QY 131 CCAAGCCAGGATCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCAGGACATCT 190

```

Db      121  CCAGCCGAGGATCTCCAGCTGGAGCAACCTCGGGCCGGGATCTCCAGCCCAAGCATCT
Qy      191  CCAGCTGGTACACCTCCAGAGCGGGGCACTCCAGAGCGGGGATCTCCAGCCCAAGCATCT
Db      181  CCAGCTGGTACACCTCCAGAGCGGGGCACTCCAGAGCGGGGATCTCCAGCCCAAGCATCT
Qy      251  CCAGCCCGGGGATCTCCGGGCTCTGGGATCACTTTCCAGAGTCTGATCCGGAGGATCACTCA
Db      241  CCAGCCCGGGGATCTCCGGGCTCTGGGATCACTTTCCAGAGTCTGATCCGGAGGATCACTCA
Qy      311  TCCGCGAGGTGAGCTCGGTGACAACTCCCAACAGAGTGTACTTTAGAGCAACA
Db      301  TCCGCGAGGTGAGCTCGGTGACAACTCCCAACAGAGTGTACTTTAGAGCAACA
Qy      371  CCAGTGGGGGGGTATACCATCGATCTCTGCGAGGTGAGAGCAAGCAAGCAAGGAGCC
Db      361  CCAGTGGGGGGGTATACCATCGATCTCTGCGAGGTGAGAGCAAGCAAGCAAGGAGCC
Qy      431  ACCAGGAGAGCCCAAGGTACAGAGCTGCGCAAGTTCACTTGGCGGAGAGCCAGAGCAG
Db      421  ACCAGGAGAGCCCAAGGTACAGAGCTGCGCAAGTTCACTTGGCGGAGAGCCAGAGCAG
Qy      491  CTACCGCTCATCGGGTGGGTGCTCTCTCTCATTTGCGCTGTGTGTTGCTCATCTCTC
Db      481  CTACCGCTCATCGGGTGGGTGCTCTCTCTCATTTGCGCTGTGTGTTGCTCATCTCTC
Qy      551  TTCCAGTTCTGGAGGGGCAACAGAGGATCAGGTACAGAGAGAGAGAGAGAGGCTGTCCC
Db      541  TTCCAGTTCTGGAGGGGCAACAGAGGATCAGGTACAGAGAGAGAGAGAGGCTGTCCC
Qy      611  AAGCAGCTGTTGCTGCTGAGCGGGGTGTGTGAGCTGCAAGCTGAAAGTGAAGAGCTGGCC
Db      601  AAGCAGCTGTTGCTGCTGAGCGGGGTGTGTGAGCTGCAAGCTGAAAGTGAAGAGCTGGCC
Qy      671  TGCGTGAAGTTTGACTGGGACAAAGTCTTGTAAATCTACTGGGTCTTCCCATCAG
Db      661  TGCGTGAAGTTTGACTGGGACAAAGTCTTGTAAATCTACTGGGTCTTCCCATCAG
Qy      731  TGCGTCCCATCTGTAGAGAGCACTGGAAATGACTCTACTAGAGAAAGAGAGCTGCGACAG
Db      721  TGCGTCCCATCTGTAGAGAGCACTGGAAATGACTCTACTAGAGAAAGAGAGCTGCGACAG
Qy      791  CTGGGTTTCGAGAGTGTCTCACCGGACAAACGAGGTGGCCCAAGGATTTTTCGCAAGAGC
Db      781  CTGGGTTTCGAGAGTGTCTCACCGGACAAACGAGGTGGCCCAAGGATTTTTCGCAAGAGC
Qy      851  TTCTCAATCTTGAATATACAACTCCACCATCCAGAGAAAGCTCCACAGGTCTGAATGCCCT
Db      841  TTCTCAATCTTGAATATACAACTCCACCATCCAGAGAAAGCTCCACAGGTCTGAATGCCCT
Qy      911  TCCGAGCGGTATATCTCCCTCCAGGTGTTCCCACTGCGGATGAGAGGCGCATGACCGGGGCG
Db      901  TCCGAGCGGTATATCTCCCTCCAGGTGTTCCCACTGCGGATGAGAGGCGCATGACCGGGGCG
Qy      971  ATCGTGGAGGGGCGCTGCGCTCGGATAGCAAGTGCGCTTGGCAAGTGAAGTCTGCACTTC
Db      961  ATCGTGGAGGGGCGCTGCGCTCGGATAGCAAGTGCGCTTGGCAAGTGAAGTCTGCACTTC
Qy      1031  GGCACCAACCACTGTGTGAGAGGCAAGCTCATTTGACGCGCCAGTGGGTGCTCACTGCGGCG
Db      1021  GGCACCAACCACTGTGTGAGAGGCAAGCTCATTTGACGCGCCAGTGGGTGCTCACTGCGGCG
Qy      1091  CACTGCTTCTTGTGACCCCGGAGAGAGGTCTGAGAGGGCTGGAAGGTGTACGCGGGCACC
Db      1081  CACTGCTTCTTGTGACCCCGGAGAGAGGTCTGAGAGGGCTGGAAGGTGTACGCGGGCACC
Qy      1151  AGCAACCTGACCAAGTTGCTGAGGAGAGCTCAATTGCGGAGATCATCATCAAGCAAT
Db      1141  AGCAACCTGACCAAGTTGCTGAGGAGAGCTCAATTGCGGAGATCATCATCAAGCAAT
Qy      1211  TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTGCCAAGCCCTGACC

```

```

Db      1201  TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTGCCAAGCCCTGACC
Qy      1271  CTGTGCGCTGACATTCACCCCTGCTGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT
Db      1261  CTGTGCGCTGACATTCACCCCTGCTGCTGCTCCCATGATGAGAGACCTTTAGCTCAAT
Qy      1331  GAGACCTGTGATTCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAAGATCCGCC
Db      1321  GAGACCTGTGATTCACAGGCTTTGGCAAGACAGAGGAGACAGATGACAAAGATCCGCC
Qy      1391  TTCTCCGGGAGGTGACAGGTTCATCTGATGACTTTCAAGAAATGCAATGACTTCTGTC
Db      1381  TTCTCCGGGAGGTGACAGGTTCATCTGATGACTTTCAAGAAATGCAATGACTTCTGTC
Qy      1451  TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGGGGGAGAGAC
Db      1441  TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGGGGGAGAGAC
Qy      1511  TCTGCCAGGAGAGACAGCGGGGGGCTCTTGTGTGTGAGAGAAACAACGCTGGTACCTG
Db      1501  TCTGCCAGGAGAGACAGCGGGGGGCTCTTGTGTGTGAGAGAAACAACGCTGGTACCTG
Qy      1571  GCAGGTGTCACAGCTGGGGGACAGAGCTGTGGCCAGAGAAACAACCTGGTGTACACC
Db      1561  GCAGGTGTCACAGCTGGGGGACAGAGCTGTGGCCAGAGAAACAACCTGGTGTACACC
Qy      1631  AAAGTGACAGAGTTCTTCCCTGGATTTACAGCAAGATGAGAGCGAGTGCATTCAGA
Db      1621  AAAGTGACAGAGTTCTTCCCTGGATTTACAGCAAGATGAGAGCGAGTGCATTCATA
Qy      1691  AAATCTTAA 1699
Db      1681  AAATCTTAA 1689

RESULT 4
DD194635
LOCUS
DEFINITION
DD194635 1689 bp DNA linear PAT 19-JAN-2006
Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
Therof.
ACCESSION
DD194635.1 GI:85634726
VERSION
JP 200518332-A/12.
KEYWORDS
unidentified
ORGANISM
unclassified
REFERENCE
1 (bases 1 to 1689)
AUTHORS
Komandla,M., Sempke,J.E., Vlaeuk,G.P., Kemp,S.J., Siev,D.V. and
Madison,E.L.
TITLE
Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
Therof
JOURNAL
Patent: JP 200518332-A 12 23-JUN-2005;
Edwin I. Madison, Joseph Edward Sempke, George P Vlaeuk, Scott Jeffrey
Kemp, Mallareddy Komandla, Daniel Vanna Siev
COMMENT
PN Homo Sapien
OS JP 200518332-A/12
PD 23-JUN-2005
PR 23-MAY-2002 JP 2002592470
PR 23-MAY-2001 US 60/293267
PI mallareddy komandla, joseph edward sempke, george p vlaeuk, pi
scott jeffrey kemp, daniel vanna siev, edwin i madison CC
FH Key Location/Qualifiers
FT CDS (1) . . . (1689)
FT /note="Nucleic acid encoding Endothelase 2-s
protein".
FEATURES
source Location/Qualifiers
1. .1689
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 96.5%; Score 1687.4; DB 2; Length 1689;

```

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 ATGAGAGGGGACACCCAGGAGATGATCTCCAGCAAGAAACCTTACGCTGAGACATCT 70
Db 1 ATGAGAGGGGACACCCAGGAGATGATCTCCAGCAAGAAACCTTACGCTGAGACATCT 60

QY 71 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 130
Db 61 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 120

QY 131 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 190
Db 121 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 180

QY 191 CCAGCTGTACCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 250
Db 181 CCAGCTGTACCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 240

QY 251 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 310
Db 241 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCGAGGATCT 300

QY 311 TCCGCAAGCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTGTAGAGACACA 370
Db 301 TCCGCAAGCTCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTGTAGAGACACA 360

QY 371 CCAGTGGGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGAGGATCT 430
Db 361 CCAGTGGGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGGCGAGGATCT 420

QY 431 ACCAGGAGAGCCAGGATCCAGGCTGGGACACCTCCAGGCGGGGATCTCCAGGAGAG 490
Db 421 ACCAGGAGAGCCAGGATCCAGGCTGGGACACCTCCAGGCGGGGATCTCCAGGAGAG 480

QY 491 CTACCGCTCAGCTGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTCTC 550
Db 481 CTACCGCTCAGCTGGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCTCTC 540

QY 551 TTCCAGTTCTGGGACAGGACACAGGATCAGGTACAGAGACAGAGGAGAGCTGTGCC 610
Db 541 TTCCAGTTCTGGGACAGGACACAGGATCAGGTACAGAGAGAGAGGAGAGCTGTGCC 600

QY 611 AAGAGCGCTGTGGCTGGGACAGGATCTGGGACAGGATCTGGGACAGGATCTGGGAC 670
Db 601 AAGAGCGCTGTGGCTGGGACAGGATCTGGGACAGGATCTGGGACAGGATCTGGGAC 660

QY 671 TGGGTGAGGTTTGGGACAGGATCTGGGACAGGATCTGGGACAGGATCTGGGACAG 730
Db 661 TGGGTGAGGTTTGGGACAGGATCTGGGACAGGATCTGGGACAGGATCTGGGACAG 720

QY 731 TGGCTTCCATCTGTAGCAGGACAGGATCTGTAGTACTCTAGAGAGAGAGCTGTCCAG 790
Db 721 TGGCTTCCATCTGTAGCAGGACAGGATCTGTAGTACTCTAGAGAGAGAGCTGTCCAG 780

QY 791 CTGGGTTTCCAGAGTGTCTACCCGAGACAGGATCTGGGACAGGATCTGGGACAGG 850
Db 781 CTGGGTTTCCAGAGTGTCTACCCGAGACAGGATCTGGGACAGGATCTGGGACAGG 840

QY 851 TTTCATCTTGAAGATCAACTCCAGCATCCAGAGAAAGCTCCAGAGTGTGAATGCCCT 910
Db 841 TTTCATCTTGAAGATCAACTCCAGCATCCAGAGAAAGCTCCAGAGTGTGAATGCCCT 900

QY 911 TCCGACGAGTATCTCCCTCCAGTGTTCCTCAGTGGGATGAGGAGCCATGACCGGAG 970
Db 901 TCCGACGAGTATCTCCCTCCAGTGTTCCTCAGTGGGATGAGGAGCCATGACCGGAG 960

QY 971 ATGCTGGAGGAGGAGCTGGGCTGGGATGAGAGGAGGCTGTGGCAAGTGTGACATCTC 1030
Db 961 ATGCTGGAGGAGGAGCTGGGCTGGGATGAGAGGAGGCTGTGGCAAGTGTGACATCTC 1020

QY 1031 GGACCAACCAATCTGTGAGGACGCTCATTTGACGCGGAGGAGTGTCTCATCTGCCGC 1090

Db 1021 GGACCAACCAATCTGTGAGGACGCTCATTTGAGCCGAGGAGTGTCTCATCTGCCGC 1080
QY 1091 CACTGCTTCTTGGGACCCGGGAGAGGCTCTGGAGGGCTGGAGGCTGACCGGGGAC 1150
Db 1081 CACTGCTTCTTGGGACCCGGGAGAGGCTCTGGAGGGCTGGAGGCTGACCGGGGAC 1140

QY 1151 AGCACTGACAGGATCTTGGGACGAGCTCCATTTGCCAGATCATCATCAGCAAT 1210
Db 1141 AGCACTGACAGGATCTTGGGACGAGCTCCATTTGCCAGATCATCATCAGCAAT 1200

QY 1211 TACACGATGAGAGAGAGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1270
Db 1201 TACACGATGAGAGAGAGAGATATGATGATGATGATGATGATGATGATGATGATGAT 1260

QY 1271 CTGTCGCTCAGATCCAGCTGGGATCTCCAGTGGGATGATGATGATGATGATGATGAT 1330
Db 1261 CTGTCGCTCAGATCCAGCTGGGATCTCCAGTGGGATGATGATGATGATGATGATGAT 1320

QY 1331 GAGACTGTGATCAGAGCTTTGGGACAGGAGACAGATGACAGATGACATCCCC 1390
Db 1321 GAGACTGTGATCAGAGCTTTGGGACAGGAGACAGATGACAGATGACATCCCC 1380

QY 1391 TTCTCCGAGAGGATGAGGATCTGATGATGATGATGATGATGATGATGATGATGAT 1450
Db 1381 TTCTCCGAGAGGATGAGGATCTGATGATGATGATGATGATGATGATGATGATGAT 1440

QY 1451 TATGACATTTACTTACCAGGATGATGATGATGATGATGATGATGATGATGATGAT 1510
Db 1441 TATGACATTTACTTACCAGGATGATGATGATGATGATGATGATGATGATGATGAT 1500

QY 1511 TCTGCGAGGAGAGACAGGAGGAGCTCTTGTCTGTGAGAGAGACAGGATGATGAT 1570
Db 1501 TCTGCGAGGAGAGACAGGAGGAGCTCTTGTCTGTGAGAGAGACAGGATGATGAT 1560

QY 1571 GCAAGTGTCAACAGCTGGGACAGGCTGTGTGAGGAGAGAGAGAGAGAGAGAGAG 1630
Db 1561 GCAAGTGTCAACAGCTGGGACAGGCTGTGTGAGGAGAGAGAGAGAGAGAGAGAG 1620

QY 1631 AAGTGAACAAGTTCTTCCCTGATTTACAGCAAGATGAGAGAGAGAGAGAGAG 1690
Db 1621 AAGTGAACAAGTTCTTCCCTGATTTACAGCAAGATGAGAGAGAGAGAGAGAG 1680

QY 1691 AAATCTTAA 1699
Db 1681 AAATCTTAA 1689

RESULT 5
AX149579 1689 bp DNA linear PAT 31-AUG-2001
LOCUS AX149579
DEFINITION Sequence 3 from Patent WO0136604.
ACCESSION AX149579
VERSION AX149579.1 GI:14348013
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1. Madison, E.L. and Ong, E.O.
Nucleic acids encoding endothelialases, endothelialases and uses
thereof
Patent: WO 0136604-A 3 25-MAY-2001;
JOURNAL CORVAS INTERNATIONAL, INC. (US)
FEATURES
source
1. .1689
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .1689
CDS
/note="unnamed protein product; DNA encoding a
transmembrane serine protease (Endothelase 2-s) protein"

```

/codon_start=1
/protein_id="CAC41220.1"
/db_xref="GI:14348014"
/translation="MERDSHGNASPARTPSAGAPASPAOSPAOTPPGRASPAOSPAO
SPACTPPGRASPAOSPAOTPPGRASPAOSPAOSPAOSPAOSPAOSPAOSPAOSPAO
RSASVTSPTPTVTVLRATPVAVPIRSSPASABATATRSPTSLPKPTWRGOKO
LPLIGCVDLILALVSLILFQPMOHTGILYKQRESCEPHARCDVDCIKSDE
LGCVRFDWDLKLTIVGSSSHQWLPICSNMNDYSEKTCQQLFESMARTTEVAHD
FANSFILRYNSTIOESLHRECEPORYISIQCSHCGLRATGRIVGALASDSKWP
OVSLHPTTHICQGLTLDIAOWLTFAHCFYTRKRVLGQWLVYAGTSLHLDPEASIA
AEIINNYDEBDYDIALMRSLKPLTSLSHIPACIPMHGPFSLNTECMTWGPGK
TRTDKTSPTLRREVQWMLDFKCNLYLVDSYLTTRPMKAGLRGRSDCGQSDSG
PLVCEONRMVLAVTSMGTGCGQKRNKPVITKVEVLPWISMESVRIKIS"

```

Query Match 96.5%; Score 1687.4; DB 2; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 11 ATGAGAGAGGACAGCCAGGAGATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
Db 1 ATGAGAGAGGACAGCCAGGAGATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 60
QY 71 CCAAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 130
Db 61 CCAAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 120
QY 131 CCAAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 190
Db 121 CCAAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 180
QY 191 CCAAGCTGTACACCTCCAGGCCGGGACATCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 250
Db 181 CCAAGCTGTACACCTCCAGGCCGGGACATCTCCAGGCCGGGACATCTCCAGGCCAGGACATCT 240
QY 251 CCAAGCCGGGATCTCCGGGCTCTGGCATCATCTTCAGAGTCTCTCATCCGGGAGGTCATCA 310
Db 241 CCAAGCCGGGATCTCCGGGCTCTGGCATCATCTTCAGAGTCTCTCATCCGGGAGGTCATCA 300
QY 311 TCCGCGAGGTCAGGCTCGGTGACAAACCTCCCAACAGAGTGTACTTGTAGAGCAACA 370
Db 301 TCCGCGAGGTCAGGCTCGGTGACAAACCTCCCAACAGAGTGTACTTGTAGAGCAACA 360
QY 371 CCAAGTGGGGGTGTACCCATCGATCATCTCTGCGAGGTGAGACCAAGCAACAGGACC 430
Db 361 CCAAGTGGGGGTGTACCCATCGATCATCTCTGCGAGGTGAGACCAAGCAACAGGACC 420
QY 431 ACCAGGAGAGACCAGAGTACAGAGCTGCGCAAGTTTCACTTGGCGGAGGACCAAGAGAG 490
Db 421 ACCAGGAGAGACCAGAGTACAGAGCTGCGCAAGTTTCACTTGGCGGAGGACCAAGAGAG 480
QY 491 CTACCGCTCATCGGAGTGTGCTCTCTCATTTGCGCTGCTGCTCATCATCTCTC 550
Db 481 CTACCGCTCATCGGAGTGTGCTCTCTCATTTGCGCTGCTGCTCATCATCTCTC 540
QY 551 TTCCAGTTCTTGAGAGGACCAAGAGATCAGGTACAAAGAGAGAGAGAGAGTGTCTCC 610
Db 541 TTCCAGTTCTTGAGAGGACCAAGAGATCAGGTACAAAGAGAGAGAGAGAGTGTCTCC 600
QY 611 AAGAGCGCTGTTGCTGTGACGGGGTGTGTGACTGCAAGCTGGAAGAGTGAAGACTGTGGC 670
Db 601 AAGAGCGCTGTTGCTGTGACGGGGTGTGTGACTGCAAGCTGGAAGAGTGAAGACTGTGGC 660
QY 671 TGCGTGAAGTTGACTGGGACAAAGTCTGTCTAAATCTACTCTGGGCTCTCCCATCAG 730
Db 661 TGCGTGAAGTTGACTGGGACAAAGTCTGTCTAAATCTACTCTGGGCTCTCCCATCAG 720
QY 731 TGCGTTCCTCATCTGTAGAGCAACTGGAATGACTCTTACTAGAGAGAGAGTGTCCAGAG 790
Db 721 TGCGTTCCTCATCTGTAGAGCAACTGGAATGACTCTTACTAGAGAGAGAGTGTCCAGAG 780
QY 791 CTGGGTTTCAGAGATGTCTCACCGGACCAAGAGTGTGCCACAGGATTTTGTCCAAAGC 850

```

```

Db 781 CTGGGTTTCAGAGATGTCTCACCGGACCAAGAGTGTGCCACAGGATTTTGTCCAAAGC 840
QY 851 TTCTCAATCTTGAGATACAACTCCACCATCCAGGAGAAAGCTCCACAGGTCTGAATGCCCT 910
Db 841 TTCTCAATCTTGAGATACAACTCCACCATCCAGGAGAAAGCTCCACAGGTCTGAATGCCCT 900
QY 911 TCCAGCGGTATATCTCCCTCAAGTGTTCACATCTGCCGACTGAGGGCCATGACCGGGCGG 970
Db 901 TCCAGCGGTATATCTCCCTCAAGTGTTCACATCTGCCGACTGAGGGCCATGACCGGGCGG 960
QY 971 ATCGTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1030
Db 961 ATCGTGGAGGGGGGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1020
QY 1031 GGCACCAACCCATCTGTGAGAGGACCGCTCATTTAGCCGCCAGTGGTCTCACTGCGGCC 1090
Db 1021 GGCACCAACCCATCTGTGAGAGGACCGCTCATTTAGCCGCCAGTGGTCTCACTGCGGCC 1080
QY 1091 CACTGCTTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAGGTGTACCGGGACAC 1150
Db 1081 CACTGCTTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAGGTGTACCGGGACAC 1140
QY 1151 AGCAACCTGACCAAGTGTGCTGAGGAGCGCTCCATTCGAGAGATCATCAAGCAAT 1210
Db 1141 AGCAACCTGACCAAGTGTGCTGAGGAGCGCTCCATTCGAGAGATCATCAAGCAAT 1200
QY 1211 TACACCGATGAGAGAGAGCACTATGATGAGCGCCCTCATGCGGCTGTCCAGGCCCTGACC 1270
Db 1201 TACACCGATGAGAGAGAGCACTATGATGAGCGCCCTCATGCGGCTGTCCAGGCCCTGACC 1260
QY 1271 CTGTGCGCTCATTCACCGCTGCTTGCCTCCCATGCAATGAGACAGACCTTTAGCTCAAT 1330
Db 1261 CTGTGCGCTCATTCACCGCTGCTTGCCTCCCATGCAATGAGACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACCTGCTGATCAACAGGCTTTGGAGACAGAGGAGACAGATGACAGACATCCGCC 1390
Db 1321 GAGACCTGCTGATCAACAGGCTTTGGAGACAGAGGAGACAGATGACAGACATCCGCC 1380
QY 1391 TTCTCCGGGAGGTGACAGTCAATCTGATGACCTTCAAGAAATGCAATGACTGATGCTC 1450
Db 1381 TTCTCCGGGAGGTGACAGTCAATCTGATGACCTTCAAGAAATGCAATGACTGATGCTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATATGTGTCTGGGACCTTGTGGGGGACAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATATGTGTCTGGGACCTTGTGGGGGACAGAGAC 1500
QY 1511 TCCGCGAGGAGAGCAAGGGGGGCTCTGTCTGTGAGAGAGCAACCGCTGGTACTCTG 1570
Db 1501 TCCGCGAGGAGAGCAAGGGGGGCTCTGTCTGTGAGAGAGCAACCGCTGGTACTCTG 1560
QY 1571 GCAAGTGTCAACAGCTGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1630
Db 1561 GCAAGTGTCAACAGCTGGGACAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1620
QY 1631 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGGTGCGATTGACA 1690
Db 1621 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGGTGCGATTGACA 1680
QY 1691 AAATCTTAA 1699
Db 1681 AAATCTTAA 1689

```

```

RESULT 6
CS081489 2380 bp DNA 1lineat PAT 18-MAY-2005
LOCUS CS081489
DEFINITION Sequence 1 from Patent WO2005040401.
ACCESSION CS081489
VERSION CS081489.1 GI:66348810
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
Bayer Healthcare AG (DE)
Location/Qualifiers
1. .2380
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES

Source

ORIGIN

Query Match 95.6%; Score 1670.4; DB 2; Length 2380;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTCAGAGCCATGAGAGGAGGACAGCCAGGGAATGCACTCCAGCAAGAACCTTCAGC 60
78 CTCAGAGCCATGAGAGGAGGACAGCCAGGGAATGCACTCCAGCAAGAACCTTCAGC 137
61 TGGAGCATCTCCAGGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 120
138 TGGAGCATCTCCAGGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 197
121 CCAGGCATCTCCAGGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 180
198 CCAGGCATCTCCAGGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 257
181 CCAGGCATCTCCAGCTGGTATACCTCCAGGCCGGGATCTCCAGGCCGGGATCTCCAGC 240
258 CCAGGCATCTCCAGCTGGTATACCTCCAGGCCGGGATCTCCAGGCCGGGATCTCCAGC 317
241 CCAGGCATCTCCAGGCCAGGATCTCCAGCTGGGATCTCCAGGCCGGGATCTCCAGC 300
318 CCAGGCATCTCCAGGCCAGGATCTCCAGCTGGGATCTCCAGGCCGGGATCTCCAGC 377
301 CAGGTCATCATCCGCGAGGATCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGT 360
378 CAGGTCATCATCCGCGAGGATCAAGCTCGGTGACAACTCCCAACAGAGTGTACTTGT 437
361 TAGAGCAACACAGTGGGGGCTGTACCATCCGATCATCTCTCCAGAGTCAAGCACAGC 420
438 TAGAGCAACACAGTGGGGGCTGTACCATCCGATCATCTCTCCAGAGTCAAGCACAGC 497
421 AACCAAGGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
498 AACCAAGGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 557
481 CCAGAGAGGATCAAGCTCATCGGGTGGGTCTCTCTCATTTGCTGCTGTGGTTCCT 540
558 CCAGAGAGGATCAAGCTCATCGGGTGGGTCTCTCTCATTTGCTGCTGTGGTTCCT 617
541 CATCATCT 600
618 CATCATCT 677
601 GAGCGTGTCCCAAGGACGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 660
678 GAGCGTGTCCCAAGGACGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 737
661 GAGCGTGTCCGCTGT 720
738 GAGCGTGTCCGCTGT 797
721 CTCCCATAGTGGCTTCCCATCTGTAGAGCAACCTGGATATCTCTCTCTCTCTCTCTCTCT 780
798 CTCCCATAGTGGCTTCCCATCTGTAGAGCAACCTGGATATCTCTCTCTCTCTCTCTCTCT 857
781 CTGCGAGGAGCTGGGTTTGAAGAGTGTCAACCGGACCAACCGAGGTTGGCCACAGGATTT 840

858 CTGCGAGGAGCTGGGTTTGAAGAGTGTCAACCGGACCAACCGAGGTTGGCCACAGGATTT 917
841 TGGCAACAGCTTTCATCTTTGAGATCAACCTCCACCATCCAGGAAGCCCTCCACAGTTC 900
918 TGGCAACAGCTTTCATCTTTGAGATCAACCTCCACCATCCAGGAAGCCCTCCACAGTTC 977
901 TGAATGCTCTTCCAGGAGTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
978 TGAATGCTCTTCCAGGAGTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1037
961 GACCGGAGGATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1038 GACCGGAGGATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
1021 TCTGCACTTGGGACCAACCACTCTGTGAGAGGACCTCTATTAACCCCAAGTGGTCT 1080
1098 TCTGCACTTGGGACCAACCACTCTGTGAGAGGACCTCTATTAACCCCAAGTGGTCT 1157
1081 CACTGCGGCCCACTGCTTCTTCTGTGAGCCGGGAGAAAGTCTCTGAGGAGGAGTGA 1140
1158 CACTGCGGCCCACTGCTTCTTCTGTGAGCCGGGAGAAAGTCTCTGAGGAGGAGTGA 1217
1141 GCGGGGACCAAGCAACCTGACCAAGTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1218 GCGGGGACCAAGCAACCTGACCAAGTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1277
1201 CAACAGCAATTAACCCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
1278 CAACAGCAATTAACCCGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
1261 GCGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
1338 GCGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1397
1321 TAGCTCAATGAGACCTGCTGATCAAGGCTTTGGAGAGCAAGGAGGAGGAGGAGGAGGAG 1380
1398 TAGCTCAATGAGACCTGCTGATCAAGGCTTTGGAGAGCAAGGAGGAGGAGGAGGAGGAG 1457
1381 GACATCCGCTTCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
1458 GACATCCGCTTCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1517
1441 CTACTGCTATGAGACCTTACCTTACCCCAAGGATGATGATGATGATGATGATGATGATGAT 1500
1518 CTACTGCTATGAGACCTTACCTTACCCCAAGGATGATGATGATGATGATGATGATGATGAT 1577
1501 GGGAGAGGATCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
1578 GGGAGAGGATCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1637
1561 CTGCTACCTGAGAGGATCAACAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
1638 CTGCTACCTGAGAGGATCAACAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1697
1621 TGTGTACCAACCAAGTGAAGAGGTTCTTCCCTGATTTACAGCAAGATGAG 1672
1698 TGTGTACCAACCAAGTGAAGAGGTTCTTCCCTGATTTACAGCAAGATGAG 1749

RESULT 7
DD200135 2393 bp DNA linear PAT 19-JAN-2006
LOCUS DD200135
DEFINITION Methods and compositions for treating cardiovascular disease.
ACCESSION DD200135.1 GI:85650814
VERSION DD200135.1
KEYWORDS JP 200511605-A/50.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2393)
AUTHORS Action,S.L., Healy,E., Suteguriano,N., Galvin,K.M., Chun,M.,
RodorigyMweil,A., Donagyu,M., Perodin,J., and Logan,T.J.
TITLE Methods and compositions for treating cardiovascular disease

JOURNAL

Patent: JP 2005516605-A 50 09-JUN-2005;

COMMENT

Millennium Pharmaceuticals Inc
OS Homosapien
PN JP 2005516605-A/50
PD 09-JUN-2005

PF 29-JAN-2003 JP 2003565410
PR 05-DEC-2002 US 60/431079, 03-JUL-2002 US 60/394130, PR
19-APR-2002 US 60/373861, 01-FEB-2002 US 60/353224, PR
05-DEC-2002 US 60/431042, 15-MAR-2002 US 60/364529, PR
04-SEP-2002 US 60/408070, 10-JUL-2002 US 60/394797, PR
21-AUG-2002 US 60/404904, 23-AUG-2002 US 60/405450, PR
06-NOV-2002 US 60/424300, 24-JUN-2002 US 60/390971, PR
12-JUN-2002 US 60/388080, 29-APR-2002 US 60/376287 PI Susan
I action, sileen healy, nancy sutagurianno, katherine m PI galvin,
PI myoung chun, amei rodorigyu-wei, marie donagyu, jacqueline PI
perodin,
PI thomas joseph logan

PI thomas joseph logan

PI thomas joseph logan

FEATURES
source Location/Qualifiers
1. 2393
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 95.6%; Score 1670.4; DB 2; Length 2393;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGAGAGACATGAGAGAGGACAGCCAGGGAATGATCTTCAGACAGAACACCTTCAGC 60
DB CTCAGAGACCATGAGAGAGGACAGCCAGGGAATGATCTTCAGACAGAACACCTTCAGC 137
QY 61 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCCGGGATCTCCAGC 120
DB TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCCGGGATCTCCAGC 197
QY 121 CCAGGCACTCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCCGGGATCTCCAGC 180
DB CCAGGCACTCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCCGGGATCTCCAGC 257
QY 181 CCAGGCACTCTCCAGCTGGTACACCTCCAGCCCGGGATCTCCAGCCCGGGATCTCCAGC 240
DB CCAGGCACTCTCCAGCTGGTACACCTCCAGCCCGGGATCTCCAGCCCGGGATCTCCAGC 317
QY 241 CCAGGCACTCTCCAGCCCGGGATCTCCAGCTGGGACACCTTCAGAGCTCTCATCCGG 300
DB CCAGGCACTCTCCAGCCCGGGATCTCCAGCTGGGACACCTTCAGAGCTCTCATCCGG 377
QY 301 CAGGTCACTCATCCGCGAGTCAAGCTCCGATCATCTTCCTGCGAGCTCAGACCAAGC 360
DB CAGGTCACTCATCCGCGAGTCAAGCTCCGATCATCTTCCTGCGAGCTCAGACCAAGC 437
QY 361 TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTTCCTGCGAGCTCAGACCAAGC 420
DB TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTTCCTGCGAGCTCAGACCAAGC 497
QY 438 TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTTCCTGCGAGCTCAGACCAAGC 497
DB TAGAGCAACACCAAGTGGGGCTGTACCCATCCGATCATCTTCCTGCGAGCTCAGACCAAGC 497
QY 421 AACAGGGCCACCAAGGAGAGCCAGGTAAGAGCTGCCAAGTTCACTTCCTGCGAGGAGG 480
DB AACAGGGCCACCAAGGAGAGCCAGGTAAGAGCTGCCAAGTTCACTTCCTGCGAGGAGG 557
QY 498 AACAGGGCCACCAAGGAGAGCCAGGTAAGAGCTGCCAAGTTCACTTCCTGCGAGGAGG 557
DB AACAGGGCCACCAAGGAGAGCCAGGTAAGAGCTGCCAAGTTCACTTCCTGCGAGGAGG 557
QY 481 CCAGAGCAGCTACCGCTCATCGGGTGGTCTCTCTCATTTGCGCTGTGTTGCT 540
DB CCAGAGCAGCTACCGCTCATCGGGTGGTCTCTCTCATTTGCGCTGTGTTGCT 617
QY 541 CATCATCTCTTCCAGATTCTGGCAGGGCCACACAGGATCAGGTACAAAGAGCAGAGGGA 600
DB CATCATCTCTTCCAGATTCTGGCAGGGCCACACAGGATCAGGTACAAAGAGCAGAGGGA 677
QY 601 GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAAGTGA 660
DB GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAAGTGA 737

QY 661 CGAGCTGGGCTGCGTGAAGTTGACTGGGACAAATCTCTGCTTAAATCTACTCTGGGTC 720
DB CGAGCTGGGCTGCGTGAAGTTGACTGGGACAAATCTCTGCTTAAATCTACTCTGGGTC 797
QY 721 CTCCATCATGAGTGGCTTCCCATCTGTAGACCAATCTGAAATGATCTCTTCAAGAGAAC 780
DB CTCCATCATGAGTGGCTTCCCATCTGTAGACCAATCTGAAATGATCTCTTCAAGAGAAC 857
QY 781 CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACAAACCGAGGTTGCCACAGGAAATT 840
DB CTGCGACAGCTGGGTTTCCAGAGTGTCAACCGGACAAACCGAGGTTGCCACAGGAAATT 917
QY 841 TGCCAAAGCTTCTCAATCTTGAATACAACTCCACCATCCAGGAAAGCTTCCACAGGTC 900
DB TGCCAAAGCTTCTCAATCTTGAATACAACTCCACCATCCAGGAAAGCTTCCACAGGTC 977
QY 901 TGAATGCCCTTCCAGCGGTATATCTCCCTCCATGATTTCCATGCGGATGAGGCCAT 960
DB TGAATGCCCTTCCAGCGGTATATCTCCCTCCATGATTTCCATGCGGATGAGGCCAT 1037
QY 961 GACCGGGCGAGTGGGAGAGGGGCGTGGCTCGGATAGCAAGTGGGCTTGGCAAGTGAG 1020
DB GACCGGGCGAGTGGGAGAGGGGCGTGGCTCGGATAGCAAGTGGGCTTGGCAAGTGAG 1097
QY 1021 TCTGCACTTGGGACCAACCAATCTGTGAGGACGCTCATTTGACGCCACAGTGGTGTCT 1080
DB TCTGCACTTGGGACCAACCAATCTGTGAGGACGCTCATTTGACGCCACAGTGGTGTCT 1157
QY 1098 TCTGCACTTGGGACCAACCAATCTGTGAGGACGCTCATTTGACGCCACAGTGGTGTCT 1157
DB TCTGCACTTGGGACCAACCAATCTGTGAGGACGCTCATTTGACGCCACAGTGGTGTCT 1157
QY 1081 CACTGCGGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTCTTGAAGGCTTGAAGGCTGA 1140
DB CACTGCGGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTCTTGAAGGCTTGAAGGCTGA 1217
QY 1141 CGGGGACCAAGCAACCTGACCAAGTGGCTGAGGAGCCCTCATTTGCCAGATCATCAT 1200
DB CGGGGACCAAGCAACCTGACCAAGTGGCTGAGGAGCCCTCATTTGCCAGATCATCAT 1277
QY 1201 CAACAGCAATTACACCGATGAGAGAGCACTATGACATGACCTCTCATTCGCGCTGTCAA 1260
DB CAACAGCAATTACACCGATGAGAGAGCACTATGACATGACCTCTCATTCGCGCTGTCAA 1337
QY 1278 CAACAGCAATTACACCGATGAGAGAGCACTATGACATGACCTCTCATTCGCGCTGTCAA 1337
DB CAACAGCAATTACACCGATGAGAGAGCACTATGACATGACCTCTCATTCGCGCTGTCAA 1337
QY 1261 GCCCTGACCTGTGCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACCTT 1320
DB GCCCTGACCTGTGCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACCTT 1397
QY 1338 GCCCTGACCTGTGCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACCTT 1397
DB GCCCTGACCTGTGCGCTCATATCCACCTGCTGCTGCCATGATGATGACAGACCTT 1397
QY 1321 TAGCCTCAATGAGACCTGTGATCACAGGCTTGGCAAGACCAAGGACAGATGACAA 1380
DB TAGCCTCAATGAGACCTGTGATCACAGGCTTGGCAAGACCAAGGACAGATGACAA 1457
QY 1398 TAGCCTCAATGAGACCTGTGATCACAGGCTTGGCAAGACCAAGGACAGATGACAA 1457
DB TAGCCTCAATGAGACCTGTGATCACAGGCTTGGCAAGACCAAGGACAGATGACAA 1457
QY 1381 GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACCTTCAAGAAATGCAATGA 1440
DB GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACCTTCAAGAAATGCAATGA 1517
QY 1458 GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACCTTCAAGAAATGCAATGA 1517
DB GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACCTTCAAGAAATGCAATGA 1517
QY 1441 CTACTTGTGCTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1500
DB CTACTTGTGCTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1577
QY 1518 CTACTTGTGCTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1577
DB CTACTTGTGCTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTG 1577
QY 1501 GGGCAGAGACTCTCTCCAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGCAGAACACG 1560
DB GGGCAGAGACTCTCTCCAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGCAGAACACG 1637
QY 1578 GGGCAGAGACTCTCTCCAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGCAGAACACG 1637
DB GGGCAGAGACTCTCTCCAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGCAGAACACG 1637
QY 1561 CTGCTACCTGGCAGAGTCAACAGCTGGGACACAGGCTGTGGCCAGAGAACAACTTGG 1620
DB CTGCTACCTGGCAGAGTCAACAGCTGGGACACAGGCTGTGGCCAGAGAACAACTTGG 1697
QY 1638 CTGCTACCTGGCAGAGTCAACAGCTGGGACACAGGCTGTGGCCAGAGAACAACTTGG 1697
DB CTGCTACCTGGCAGAGTCAACAGCTGGGACACAGGCTGTGGCCAGAGAACAACTTGG 1697
QY 1621 TGTGTACACCAAGTGAAGAGTCTTCTCTGATTTACAGCAAGTGGAG 1672
DB TGTGTACACCAAGTGAAGAGTCTTCTCTGATTTACAGCAAGTGGAG 1749
QY 601 GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAAGTGA 660
DB GAGCTGTCCCAAGACGCTGTTCCTGTGAGAGGGGTGGTGAATGCAAGCTGAAGAAGTGA 737

RESULT 8
AB048796
LOCUS

AB048796

2393 bp

mRNA

linear

PRI 22-MAR-2001

DEFINITION Homo sapiens mspl mRNA for membrane-type mosaic serine protease, complete cds.

ACCESSION AB048796

VERSION AB048796.1 GI:13429969

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Kim,D.R., Sharmin,S., Inoue,M. and Kido,H.

TITLE 1 (sites)

JOURNAL Cloning and expression of novel mosaic serine proteases with and without a transmembrane domain from human lung

PUBMED Biochim. Biophys. Acta 1518 (1-2), 204-209 (2001)

REFERENCE 11267681

AUTHORS 2 (bases 1 to 2393)

TITLE Direct Submission

JOURNAL Submitted (12-SEP-2000) Hiroshi Kido, The University of Tokushima, Division of Enzyme Chemistry, Institute for Enzyme Research, Kuramoto-cho, Tokushima 770, Japan (E-mail:kido@ier.tokushima-u.ac.jp, Tel:81-88-633-7424, Fax:81-88-633-7425)

FEATURES

source

location/Qualifiers

1..2393

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="lung"

/clone_lib="Marathon Ready cDNA (Clontech)"

1..2393

/gene="mspl"

1..87

/gene="mspl"

88..1833

/gene="mspl"

/codon_start=1

/transl_except=(pos:862..864,aa:Arg)

/transl_except=(pos:979..981,aa:His)

/transl_except=(pos:1573..1575,aa:His)

/product="membrane-type mosaic serine protease"

/protein_id="BAB39741.1"

/db_xref="GI:13429970"

/translation="MERDSHGNAIPARTPSAGSPAOPASPGATPRGRASPAAQSPAAQ SPATPGRASPAQAPGATPRGRASPAQSPARASPAQSPARASPALAISRSSGGSSA RSASVTPSPTRVYLVRATPVGAIVTRKSPARSAPATATRPSTSLPKRTREGQQQL PLPLGVLLIALVALSVLLEFQWQHGHAKRQRSCPHAVRCGVNDCKLKSDSE LGCFRPFMDKLILKIYSGSSHQMPLICSMNWDSESKTCLGFESAHRTTEVAHD PANFSFLRYNSTIOESLSHRHCPSORI SLQCHCGLRAMVTGRI VEGALASDEKPM OVSLAFGTHTH CGGTLIDAOVLDAHOCPFTVREKLVGMVVATGSINHLQLPASIS I AEIILNSVDEEDDYDIAIMRLSKPLTSLSHITPACLPMHGOFPSLNWCMLTGPKK TRFDIDKTSPIPREVQNULDPFKCNDLYLDVLTLTRMCAGIHNGRDS CQGDSGG PLVEEQNNRWYLAGVSWGTGCGQRNKRPKVTTKEVLPMWTYSMEVRSLOODTAPSR LGTSSGGPGGAPRV"

1834..2393

/gene="mspl"

2343..2348

/gene="mspl"

ORIGIN

polyA_signal、

Query Match 95.6%; Score 1670.4; DB 5; Length 2393;

Best local similarity 99.9%; Pred.No. 0;

Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy      1 CTCGAGACCTTCGAGAGGAGGAGCACGCACGAGGAATGATCTCCACGAGAACAACCTTAGC 60
Db      78 CTCAGAGACCATTGGAGAGGAGGACGCCAGGAGATGATCTCCAGCAGAACACTTCAGC 137
Oy     61 TGAAGCATCTCCAGCCCGACGAGCATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCAGC 120
Db    138 TGAGAGCATCTCCAGCCCGACGAGCATCTCCAGCTGGGACACCTCCAGCCCGGAGCATCTCAGC 197

```

QY	121	CCAGGCATCTCCAGACCCAGGACATTCACGCTGGGACACACTCCGGGCGGGGACATTCACG	180
Db	198	CCAGGCATCTCCAGACCCAGGACATTCAGCTGGGACACTCCGGGCGGGGACATTCAGC	257
QY	181	CCAGGCATCTCCAGCTGGTACACTCCAGGCGGGGACATTCACAGCGGGGACATTCAGC	240
Db	258	CCAGGCATCTCCAGCTGGTACACTCCAGGCGGGGACATTCACAGCGGGGACATTCAGC	317
QY	241	CCAGGCATCTCCAGACCCGGGACATTCGGGCTTGGACATCACTTTCAGGTCTCATCCG	300
Db	318	CCAGGCATCTCCAGACCCGGGACATTCGGGCTTGGACATCACTTTCAGGTCTCATCCG	377
QY	301	CAGGTCAATCATCCGCCAGGTCAAGCTCTGGTGAACAACCTCCCAACAGAGTATCTTGT	360
Db	378	CAGGTCAATCATCCGCCAGGTCAAGCTCTGGTGAACAACCTCCCAACAGAGTATCTTGT	437
QY	361	TAGAGCAACACAACTGGGGGCTGAACCATTCGATCAATCTCTGSCAGAGTACAGACAG	420
Db	438	TAGAGCAACACAACTGGGGGCTGAACCATTCGATCAATCTCTGSCAGAGTACAGACAG	497
QY	421	AACGAGGGCCACACAGGAGAGCCAGGATGACGAGCTGCCAAATTTCACTTGGCGGAGG	480
Db	498	AACGAGGGCCACACAGGAGAGCCAGGATGACGAGCTGCCAAATTTCACTTGGCGGAGG	557
QY	481	CCAGAGAGACTACCGCTCATCGGAGTGGTCTCTCTCTCATTTGCCCTGTGTTTGGCT	540
Db	558	CCAGAGAGACTACCGCTCATCGGAGTGGTCTCTCTCTCATTTGCCCTGTGTTTGGCT	617
QY	541	CATCATCTCTCTCCAGATTCGGACAGGGCCACACAGGATACAGATACAAAGAGACAGAGGA	600
Db	618	CATCATCTCTCTCCAGATTCGGACAGGGCCACACAGGATACAGGACAAAGAGACAGAGGA	677
QY	601	GAGCTGTCCCAAGACACTGTTCCACTGTGAACGGGGTGTGGATCTGCAAGCTGAAGATGA	660
Db	678	GAGCTGTCCCAAGACACTGTTCCACTGTGAACGGGGTGTGGATCTGCAAGCTGAAGATGA	737
QY	661	CGAGCTGGGCTGCGTGAAGTTGACTGGGACAAAGTCTCTGTTAAATCTACTCTGGGCTC	720
Db	738	CGAGCTGGGCTGCGTGAAGTTGACTGGGACAAAGTCTCTGTTAAATCTACTCTGGGCTC	797
QY	721	CTCCCATACAGTGGCTTCCCATCTGTACAGACAACTGGAAATGAATCTCTCACTCAGAGAAAC	780
Db	798	CTCCCATACAGTGGCTTCCCATCTGTACAGACAACTGGAAATGAATCTCTCACTCAGAGAAAC	857
QY	781	CTGCAGAGACTGGGTTTTCGAGATGTGCACCCGACAAACGAGGTTGCCACAGGATTT	840
Db	858	CTGCAGAGACTGGGTTTTCGAGATGTGCACCCGACAAACGAGGTTGCCACAGGATTT	917
QY	841	TGCCAACAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCTTCCACAGGTC	900
Db	918	TGCCAACAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCTTCCACAGGTC	977
QY	901	TGAATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCATGCGGACTGAGGGCCAT	960
Db	978	TGAATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCATGCGGACTGAGGGCCAT	1037
QY	961	GACCGGGCGGATCGTGGGAGGGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG	1020
Db	1038	GACCGGGCGGATCGTGGGAGGGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG	1097
QY	1021	TCTGCACCTTCGGGACCAACCCACATCTGTGAGAGCAGCTCATTTGACGCCACAGTGGTCT	1080
Db	1098	TCTGCACCTTCGGGACCAACCCACATCTGTGAGAGCAGCTCATTTGACGCCACAGTGGTCT	1157
QY	1081	CACGTCCGCCCACTGTCTTCTTGAGACCCGGGAGAAAGTCTCTGAGGGGCTTGGAAAGTGTAT	1140
Db	1158	CACGTCCGCCCACTGTCTTCTTGAGACCCGGGAGAAAGTCTCTGAGGGGCTTGGAAAGTGTAT	1217
QY	1141	CGGGGGACACGACAACTGACACAGATTGCGCTGAGAGGAGCTCCATTGSCGAGATCATCAT	1200
Db	1218	CGGGGGACACGACAACTGACACAGATTGCGCTGAGAGGAGCTCCATTGSCGAGATCATCAT	1277
QY	1201	CAACGCAATTACACGATGAGAGAGCACTATGACATCGCCCTCAATGGCGTGTCAAA	1260

Db	799	CTCCATCAGTGGCTTCCCATCTGTAGAGCAATGGAAATGACTCTCTAATCAGAGAAAGAC	858
Qy	781	CTGCGACAGCTGGGGTTTCCGAAGTGCTCACCGGACAACGAGGTTGCCACAGGATTT	840
Db	859	CTGCGACAGCTGGGGTTTCCGAAGTGCTCACCGGACAACGAGGTTGCCACAGGATTT	918
Qy	841	TGCGAACAAGCTTCTCAATCTTGAATACATCTCCACATCCAGAAAGGCTCCACAGCTC	900
Db	919	TGCGAACAAGCTTCTCAATCTTGAATACATCTCCACATCCAGAAAGGCTCCACAGCTC	978
Qy	901	TGAATGCCCTTCCAGCGGGTATATCTCCCTCAGAGTTCCACTGCGGACTGAGGGCCAT	960
Db	979	TGAATGCCCTTCCAGCGGGTATATCTCCCTCAGAGTTCCACTGCGGACTGAGGGCCAT	1038
Qy	961	GACCGGCGGATCTGTGGAGGGGGCGCTGGCTCGGATAGCAATGGTCCCTTGGCAATGAG	1020
Db	1039	GACCGGCGGATCTGTGGAGGGGGCGCTGGCTCGGATAGCAATGGTCCCTTGGCAATGAG	1098
Qy	1021	TTCGCACTTTGGGGAACAACCAATCTGTGAGAGGACGCTCATTTACGCCCAAGTGGTGGCT	1080
Db	1099	TTCGCACTTTGGGGAACAACCAATCTGTGAGAGGACGCTCATTTACGCCCAAGTGGTGGCT	1158
Qy	1081	CATGCGCGCCCACTGCTTCTTCTGATACCCGGGAGAAAGGCTCTGAGAGGCTGGAAGTGA	1140
Db	1159	CATGCGCGCCCACTGCTTCTTCTGATACCCGGGAGAAAGGCTCTGAGAGGCTGGAAGTGA	1218
Qy	1141	CGCGGGACCAAGCAACCTGCACACAGTTGCTTGAAGCAAGCTTCATTTGCCAGATCATAT	1200
Db	1219	CGCGGGACCAAGCAACCTGCACACAGTTGCTTGAAGCAAGCTTCATTTGCCAGATCATAT	1278
Qy	1201	CAACAGCAATTAACCCGATGAGAGAGACATATGACATCGCCTCATGCGGCTGTCCAA	1260
Db	1279	CAACAGCAATTAACCCGATGAGAGAGACATATGACATCGCCTCATGCGGCTGTCCAA	1338
Qy	1261	GCCCCCTAACCTGTGCGGCTCAATCCACCCCTGCTTCCCTCCCAATGCAATGACAAGCTT	1320
Db	1339	GCCCCCTAACCTGTGCGGCTCAATCCACCCCTGCTTCCCTCCCAATGCAATGACAAGCTT	1398
Qy	1321	TAGCCTCAATGAGACTGTGCTGATCACAGGCTTTGGCAAGACACAGGAGACAGATGACAA	1380
Db	1399	TAGCCTCAATGAGACTGTGCTGATCACAGGCTTTGGCAAGACACAGGAGACAGATGACAA	1458
Qy	1381	GACATCCCCCTTCTCCGGGAGGAGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGA	1440
Db	1459	GACATCCCCCTTCTCCGGGAGGAGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGA	1518
Qy	1441	CTACTTGGTCTATGACAGTTACTTTACCCCAAGATGATGTGTGCTGGGGACCTTTCGGG	1500
Db	1519	CTACTTGGTCTATGACAGTTACTTTACCCCAAGATGATGTGTGCTGGGGACCTTTCGGG	1578
Qy	1501	GGGCGAGAGACTCTCGCCAGGGGAGACAGGGGGGGGCTTTGTCTGTGAGCAAAACCG	1566
Db	1579	GGGCGAGAGACTCTCGCCAGGGGAGACAGGGGGGGGCTTTGTCTGTGAGCAAAACCG	1638
Qy	1561	CTGCTACCTGGCAGAGTGTCAACAGCTGGGGGACAGCTGTGGCCAGAGAAAACAACTGG	1620
Db	1639	CTGCTACCTGGCAGAGTGTCAACAGCTGGGGGACAGCTGTGGCCAGAGAAAACAACTGG	1698
Qy	1621	TGTGTACACCAAAATGACAGAAATTTCTTCTCGATTTACACGCAAGATGAGAGAGAGGT	1680
Db	1699	TGTGTACACCAAAATGACAGAAATTTCTTCTCGATTTACACGCAAGATGAGAGAGAGGT	1758
Qy	1681	GGCAATTCAGAA 1691	
Db	1759	GGGTTGAAAAA 1769	
RESULT 10			
CS075313			
LOCUS			
DEFINITION Sequence 1 from Patent WO2005019270.			
ACCESSION CS075313.1 GI:63091664			
VERSION			
2067 bp DNA linear PAT 05-WAY-2005			

[illegible]

```

Db      481  CTACCGCTCATCGGGTGGTCTCTCTCATGCGCCCTGGTGTTCCTCATCTCTC
Qy      551  TTCCAGTTCTGGCAGGGCCACACAGGATCAGGTACAAAGAGAGAGAGAGCTGTCC
Db      541  TTCCAGTTCTGGCAGGGCCACACAGGATCAGGTACAAAGAGAGAGAGAGCTGTCC
Qy      611  AAGCAGCTGTCTGCTGTGACGGGGTGGTGAAGTGAAGCTGAAGAGTGAAGCTGGG
Db      601  AAGCAGCTGTCTGCTGTGACGGGGTGGTGAAGTGAAGAGTGAAGAGTGGG
Qy      671  TGCGTGAAGTTTGAAGTGAAGAGTCTGTCTTAAATCTAAGTCTGTCTCTCCATCAG
Db      661  TGCGTGAAGTTTGAAGTGAAGAGTCTGTCTTAAATCTAAGTCTGTCTCTCCATCAG
Qy      731  TGCGTTCCTCATCTGTAGACAGCACTGGAATGACTCTCTAGAGAAAGCTGCGACAG
Db      721  TGCGTTCCTCATCTGTAGACAGCACTGGAATGACTCTCTAGAGAAAGCTGCGACAG
Qy      791  CTGGGTTTGAAGAGTCTCACCGGACAAACGAGGTTGCCCAAGGATTTTGGCAACAGC
Db      781  CTGGGTTTGAAGAGTCTCACCGGACAAACGAGGTTGCCCAAGGATTTTGGCAACAGC
Qy      851  TTCTCAATCTTGAGATTAACATCCACCATCCAGAAAGCCCTCCACAGTCTGAATGCCCT
Db      841  TTCTCAATCTTGAGATTAACATCCACCATCCAGAAAGCCCTCCACAGTCTGAATGCCCT
Qy      911  TCCGACGCGTATATCTCCCTCCAGTGTCCCACTGCGAGTGAAGGACCATGCGGCGG
Db      901  TCCGACGCGTATATCTCCCTCCAGTGTCCCACTGCGAGTGAAGGACCATGCGGCGG
Qy      971  ATCGTGGAGGGGCGCTGCGCTCGATAGAAAGTGGCCCTTGGCAAGTGTCTGCACTTC
Db      961  ATCGTGGAGGGGCGCTGCGCTCGATAGAAAGTGGCCCTTGGCAAGTGTCTGCACTTC
Qy      1031  GGACACACCCATCTGTGAGAGGACGCTCATTAAGCCGCGAGTGGTCTCATCTGCCGCC
Db      1021  GGACACACCCATCTGTGAGAGGACGCTCATTAAGCCGCGAGTGGTCTCATCTGCCGCC
Qy      1091  CACTGCTTCTTCTGTCACCGGGAGAGAGTCTGAGAGGGCTGGAAGTGTACGCGGACCC
Db      1081  CACTGCTTCTTCTGTCACCGGGAGAGAGTCTGAGAGGGCTGGAAGTGTACGCGGACCC
Qy      1151  AGCAACCTGACCAAGTTGCTTGAAGCAGCCTTCATTCGCAAGATCATCAACAGCAAT
Db      1141  AGCAACCTGACCAAGTTGCTTGAAGCAGCCTTCATTCGCAAGATCATCAACAGCAAT
Qy      1211  TAGACCGATGAGAGAGAGAGTATAGATGAGCCCTCATAGCGGCTGTCCAAAGCCCTGACC
Db      1201  TAGACCGATGAGAGAGAGAGTATAGATGAGCCCTCATAGCGGCTGTCCAAAGCCCTGACC
Qy      1271  CTGTCCGCTCATTCACACCTGCTTGTGCTCCCATGATGAGACAGACCTTTAGCTCAAT
Db      1261  CTGTCCGCTCATTCACACCTGCTTGTGCTCCCATGATGAGACAGACCTTTAGCTCAAT
Qy      1331  GAGACCTGCTGATCAACAGGCTTTTGGCAAGACAGAGAGACAGATGACAAAGATCCCCC
Db      1321  GAGACCTGCTGATCAACAGGCTTTTGGCAAGACAGAGAGACAGATGACAAAGATCCCCC
Qy      1391  TTCTCCGAGGAGTGAAGTGAATCTCATGAGCTTCAAGAAATGCAATGATTAATCTGCTC
Db      1381  TTCTCCGAGGAGTGAAGTGAATCTCATGAGCTTCAAGAAATGCAATGATTAATCTGCTC
Qy      1451  TATGACAGTTACTTACCTCCCAAGATGATGTGCTGAGGAGCTTTCGTTGGGGGAGAGAC
Db      1441  TATGACAGTTACTTACCTCCCAAGATGATGTGCTGAGGAGCTTTCGTTGGGGGAGAGAC
Qy      1511  TCCTGCAAGGAGAGACAGCGGGGGGCTCTTGTGTGTGAGCAGAAACAACCGTGTACTGTG
Db      1501  TCCTGCAAGGAGAGACAGCGGGGGGCTCTTGTGTGTGAGCAGAAACAACCGTGTACTGTG
Qy      1571  GCAGGTGTACACACTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGTGTGTACTACC

```

```

Db      1561  GCAGGTGTACACACTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGTGTACTACC
Qy      1631  AAGTGAAGAGAGTCTTCCCTGGATTATACAGAAATGAGAGAC 1675
Db      1621  AAGTGAAGAGAGTCTTCCCTGGATTATACAGAAATGAGAGAC 1665

RESULT 11
DD194636
LOCUS      DD194636
DEFINITION  DD194636 2067 bp DNA linear PAT 19-JAN-2006
            Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
            thereof.
ACCESSION  DD194636
VERSION    DD194636.1 GI:85634727
KEYWORDS   JP 200518332-A/13.
SOURCE     unidentified
ORGANISM   unidentified
            unclassified sequences.
REFERENCE  1 (bases 1 to 2067)
            Komanda,M., Sempke,J.E., Vlaeuk,G.P., Kemp,S.J., Slev,D.V. and
            Madison,B.L.
            Conjugates Activated By Cell Surface Proteases and Therapeutic Uses
            thereof
            Patent: JP 200518332-A 13 23-JUN-2005:
            Edwin L. Madison, Joseph Edward Sempke, George P Vlaeuk, Scott Jeffrey
            Kemp, Mallareddy Komanda, Daniel Vanna Slev
            OS Homo Sapien
            PN JP 200518332-A/13
            PD 23-JUN-2005
            PF 23-MAY-2002 JP 2002592470
            PR 23-MAY-2001 US 60/293267
            PI mallareddy komanda, joseph edward sempke, george p vlaeuk, pi
            scott jeffrey kemp, daniel vanna slev, edwin l madison cc
            FH Key Location/Qualifiers
            FT CDS (1) (2067)
               /note="Nucleic acid encoding (endothelase FT
               2-L protein".
FEATURES
    source          location/Qualifiers
                     1..2067
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN
Query Match 95.2%; Score 1663.4; DB 2; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

11 ATGAGAGGGACAGCCACGGGATGCAATCTCCAGCAAGAACCTTCAAGTGAAGCATCT 70
1 ATGAGAGGGACAGCCACGGGATGCAATCTCCAGCAAGAACCTTCAAGTGAAGCATCT 60
71 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGGATCTCCAGCCAGGACATCT 130
61 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGGATCTCCAGCCAGGACATCT 120
131 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGGATCTCCAGCCAGGACATCT 190
121 CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGGATCTCCAGCCAGGACATCT 180
191 CCAGCTGTATCACTCCAGGCTGGGACATCTCCAGGCTGGGATCTCCAGGCTGGGATCT 250
181 CCAGCTGTATCACTCCAGGCTGGGACATCTCCAGGCTGGGATCTCCAGGCTGGGATCT 240
251 CCAGCCGAGGATCTCCAGCTGGGACATCTCCAGGCTGGGATCTCCAGGCTGGGATCTCA 310
241 CCAGCCGAGGATCTCCAGCTGGGACATCTCCAGGCTGGGATCTCCAGGCTGGGATCTCA 300
311 TCCGCAAGTCTAGCTCGTGTGACAACTCCCAACAGAGTGTACTTGTATTAGAGACA 370
301 TCCGCAAGTCTAGCTCGTGTGACAACTCCCAACAGAGTGTACTTGTATTAGAGACA 360
371 CCAAGTGGGGGCTGTACCAATCCATCATCTCTGCGCAGGTGACGACAGCAACGAGGCC 430

```

Db	361	CCAGTGGGGGCTGTACCCATCCGATCATCTCTCCAGGTAGCAACGACACGAGGCC	420
Oy	431	ACCGAGGAGAGCCCGAGGTACGAGCCTGCCCAAGTTCACTGGCGGAGAGGCCAGAAAGCAG	490
Db	421	ACCGAGGAGAGCCCGAGGTACGAGCCTGCCCAAGTTCACTGGCGGAGAGGCCAGAAAGCAG	480
Oy	491	CTACCGCTCATCGGGTGGCTGTCTCTCTCAATTGCTCGTGTGTTTGGCTCATCATCTCTC	550
Db	481	CTACCGCTCATCGGGTGGCTGTCTCTCTCAATTGCTCGTGTGTTTGGCTCATCATCTCTC	540
Oy	551	TTCCAGTTCTGGCAGGGCCACAGGGATCTAGGTACAGAGACAGAGGAGAGAGCTGTGCC	610
Db	541	TTCCAGTTCTGGCAGGGCCACAGGGATCTAGGTACAGAGACAGAGGAGAGAGCTGTGCC	600
Oy	611	AAGCACGCTGTTCGCTGTGACGGGGTGGTGTGACTGTCAAGCTGAAGAGTGAAGAGCTGGGC	670
Db	601	AAGCACGCTGTTCGCTGTGACGGGGTGGTGTGACTGTCAAGCTGAAGAGTGAAGAGCTGGGC	660
Oy	671	TGCGTAGAGTTTGACTGGGACAAAGTCTGTGCTTAAATCTACTGTGGTCTCTCCCATAG	730
Db	661	TGCGTAGAGTTTGACTGGGACAAAGTCTGTGCTTAAATCTACTGTGGTCTCTCCCATAG	720
Oy	731	TGGCTTCCCATCTGTAGACGCACTGGAAATGACTCTACTAGAGAGAACTTGCCACGAG	790
Db	721	TGGCTTCCCATCTGTAGACGCACTGGAAATGACTCTACTAGAGAGAACTTGCCACGAG	780
Oy	791	CTGGGTTTCGAGAGTGTCAACCGGACAAACGAGGTGGCCACAGGAAATTTTGGCAACGC	850
Db	781	CTGGGTTTCGAGAGTGTCAACCGGACAAACGAGGTGGCCACAGGAAATTTTGGCAACGC	840
Oy	851	TTCTCAATCTTGAAGATACAACTCCACATCCAGGAAAGCTTCCACAGGTCTGAATGCTCT	910
Db	841	TTCTCAATCTTGAAGATACAACTCCACATCCAGGAAAGCTTCCACAGGTCTGAATGCTCT	900
Oy	911	TCCGACGGGTATATCTCCTTCCAGTGTTCACCTGCGGACTGAGGGCCATACCGGGGCG	970
Db	901	TCCGACGGGTATATCTCCTTCCAGTGTTCACCTGCGGACTGAGGGCCATACCGGGGCG	960
Oy	971	ATCGGGGAGGGGGCGCTGGCCCTGGGATGAGAAAGGCCCTTGGCAAGTAGAGCTGCACCTTC	1030
Db	961	ATCGGGGAGGGGGCGCTGGCCCTGGGATGAGAAAGGCCCTTGGCAAGTAGAGCTGCACCTTC	1020
Oy	1031	GGCACCAACCCACATCTGTGAGAGGACGCTCATTTGACGCCACAGTGGGTCTCATGCGGCC	1090
Db	1021	GGCACCAACCCACATCTGTGAGAGGACGCTCATTTGACGCCACAGTGGGTCTCATGCGGCC	1080
Oy	1091	CACGTGCTTCTTGTGACCCGGGAGAGAGTCTGTGAGGGCTGGAAGGTGTACGCGGGCAC	1150
Db	1081	CACGTGCTTCTTGTGACCCGGGAGAGAGTCTGTGAGGGCTGGAAGGTGTGTACGCGGGCAC	1140
Oy	1151	AGCAACCTGCACAGATTGGCTTGAGGACACCTCCATTGGCCGAGATCATATCAACAGCAAT	1210
Db	1141	AGCAACCTGCACAGATTGGCTTGAGGACACCTCCATTGGCCGAGATCATATCAACAGCAAT	1200
Oy	1211	TACACCGATGAGAGAGAGCATATATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGAC	1270
Db	1201	TACACCGATGAGAGAGAGCATATATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGAC	1260
Oy	1271	CTGTCCGCTCAATCCACCCCTGTGCTTCCCATGACATGACAGACCTTTAGCTCAAT	1330
Db	1261	CTGTCCGCTCAATCCACCCCTGTGCTTCCCATGACATGACAGACCTTTAGCTCAAT	1320
Oy	1331	GAGACCTGCTGATACACAGGCTTTTGGCAAGACAGGAGACAGATGACAAACAATCCCC	1390
Db	1321	GAGACCTGCTGATACACAGGCTTTTGGCAAGACAGGAGACAGATGACAAACAATCCCC	1380
Oy	1391	TTCTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTACTTTGGTTC	1450
Db	1381	TTCTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTACTTTGGTTC	1440
Oy	1451	TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGACCTTGTGGGGGCAAGAC	1510

DB	1441	TATACAGTTACCTTACCACAGATATATGTGTCTGGGACCTTGTGGGGGACAGAC	1500
QY	1511	TCCTGCCAGGAGACACGCGGGGGCTCTTGTCTGTGAGACAAACCGCTGGTACTGT	1570
DB	1501	TCCTGCCAGGAGACACGCGGGGGCTCTTGTCTGTGAGACAAACCGCTGGTACTGT	1560
QY	1571	GCAGTGTCCACACTGTGGGACACAGCGCTGTGGCCAGAGAAACAACTGGTGTACAC	1630
DB	1561	GCAGTGTCCACACTGTGGGACACAGCGCTGTGGCCAGAGAAACAACTGGTGTACACC	1620
QY	1631	AAAGTGACAGAAAGTTCTTCCCTGGATTATTAAGACAGATGAGAGC	1675
DB	1621	AAAGTGACAGAAAGTTCTTCCCTGGATTATTAAGACAGATGAGAGAC	1665
RESULT 12			
LOCUS	AXI49581	2067 bp	DNA
DEFINITION	Sequence 5 from Patent WO0136604.		
ACCESSION	AXI49581		
VERSION	AXI49581.1		
KEYWORDS	GI:14348015		
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
	1. Madison, E.L. and Ong, E.O.		
	Nucleic acids encoding endothelialases, endothelialases and uses		
	thereof		
	Patent: WO 0136604-A 5 25-MAY-2001;		
JOURNAL	CORVAS INTERNATIONAL, INC. (US)		
FEATURES	location/Qualifiers		
source	1..2067		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	1..2067		
	/note="unnamed protein product; DNA sequence encoding a		
	transmembrane serine protease (endothelase-L) protein"		
	/codon_start=1		
	/protein_id="CAC41221.1"		
	/db_xref="GI:14348016"		
	/translation="MERDSHGNSPARTSPAGSPAOASPACTPPGRASPAPASPAO		
	SPGTPPGRASPAPASPAAGPPGRASPAPASPAOASPARAPALASISGRSSSA		
	RSASVTSPPRVYLVRATPGVAIPISSPARAPATPAPRESGTSIPKTRMEGKO		
	LPLGLFLLIALVALVSLILPFGVGHGTGRIVEQRESCPKRAVRCDGVYDCKLSE		
	LGCRFPMDKSLTKIYESSSHOWLPICSSWNDSYSEKTCQQLGFESAHRTTEVAHD		
	FANFSILRYNSTIOESLHRSCEPSQRYISLCSHGLRMTGRIVGALASDSKNPW		
	QVSLHFGTTHICGGTILDAQWVLTAAHCFVTESEKVLGKRVAGTSLHQLDEASII		
	ARILINTNSTDDEEDVDIALMRLSKIPFLTSVAHHPACLPHMGOTFSLINETCWTGK		
	TRTFDDTSPFLREVOYNLIDPFKNDYLVLYDSILTRPRMKCAEDLKGRRSCGDSG		
	PLVEQNNRWYLAGVTSWGTCGQNRNPGVYTVTEVLPIWIKEMNRARVEAMTY		
	RPFGQLGRCSPSRIPLCKVAMPENVSVAEDFVIVFKHLCMGIRSSWPALFV		
	LVFLIFFLILLSFLKNTSDSIILTLTFVATVMTLPENVSFPFLKIFQNTLTIK		
	EYKIK"		
ORIGIN			
Query Match	95.2%; Score 1663.4; DB 2; Length 2067;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1664; Conservative %	0; Mismatches 1; Indels 0; Gaps 0;		
QY	11	ATGAGAGGGAACACCCACGGGAATGCACTCCAGCAAGAACACCTTCAGCTGAGCACT	70
DB	1	ATGAGAGGGAACACCCACGGGAATGCACTCCAGCAAGAACACCTTCAGCTGAGCACT	60
QY	71	CCAGCCAGGCACTCCAGCTGGGACACCTCCAGCGCGGCACTCCAGCCCAAGCACT	130
DB	61	CCAGCCAGGCACTCCAGCTGGGACACCTCCAGCGCGGCACTCCAGCCCAAGCACT	120
QY	131	CCAGCCAGGCACTCCAGCTGGGACACCTCCGCGCGGCACTCCAGCCCAAGCACT	190

Db 121 CCAGCCAGGAGATCTCCAGCTGGGAGACCTCCGGCCGGGATCTCCAGCCAGGATCT 180
 Qy 191 CCAGCTGGTACACCTCCAGCCGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 250
 Db 181 CCAGCTGGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGATCT 240
 Qy 251 CCAGCCGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCACTCCGGAGGATCA 310
 Db 241 CCAGCCGGGATCTCCGGCTCTGGGATCACTTTCCAGGCTCTCACTCCGGAGGATCA 300
 Qy 311 TCCGCGAGGATGAGCTCGGATCAACCTCCCAACGAGATGATCTTTGATGAGCA 370
 Db 301 TCCGCGAGGATGAGCTCGGATCAACCTCCCAACGAGATGATCTTTGATGAGCA 360
 Qy 371 CCAGTGGGGGATGATCCGATCATCTCTGCGAGGATGAGCAACGAGGATCT 430
 Db 361 CCAGTGGGGGATGATCCGATCATCTCTGCGAGGATGAGCAACGAGGATCT 420
 Qy 431 ACCAGGAGAGAGCCAGGATGAGAGCTGCGCAAGTTCACTGGCGGGAGGCGCAGAG 490
 Db 421 ACCAGGAGAGAGCCAGGATGAGAGCTGCGCAAGTTCACTGGCGGGAGGCGCAGAG 480
 Qy 491 CTACCGCTCATCGGGTGGGTCTCTCTCATTTGGCTGGTGGTGGTGGTGGTGGTGGT 550
 Db 481 CTACCGCTCATCGGGTGGGTCTCTCTCATTTGGCTGGTGGTGGTGGTGGTGGTGGT 540
 Qy 551 TTCCAGTTCTGGAGGAGCCAGAGGATGAGATGAGAGAGAGAGAGAGAGAGAGAG 610
 Db 541 TTCCAGTTCTGGAGGAGCCAGAGGATGAGATGAGAGAGAGAGAGAGAGAGAGAG 600
 Qy 611 AAGCAGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 670
 Db 601 AAGCAGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
 Qy 671 TGGGAGAGGTTGAGTGGGAGCAAGCTCTGCTTAAATCTAATCTGAGGCTCTCCAT 730
 Db 661 TGGGAGAGGTTGAGTGGGAGCAAGCTCTGCTTAAATCTAATCTGAGGCTCTCCAT 720
 Qy 731 TGGCTTCCCATCTGAGAGCAACTGGAATGATCTCTAAGAGAGAGAGAGAGAGAG 790
 Db 721 TGGCTTCCCATCTGAGAGCAACTGGAATGATCTCTAAGAGAGAGAGAGAGAGAG 780
 Qy 791 CTGGGTTTGGAGAGTCTCACCGGACCAAGGTTGCCACAGGATTTTTCACAGC 850
 Db 781 CTGGGTTTGGAGAGTCTCACCGGACCAAGGTTGCCACAGGATTTTTCACAGC 840
 Qy 851 TTCTCAATCTTGAATTAACAATCTCCAGCATCTCAAGAAAGCTCCACAGGTTGAG 910
 Db 841 TTCTCAATCTTGAATTAACAATCTCCAGCATCTCAAGAAAGCTCCACAGGTTGAG 900
 Qy 911 TCCGAGCGGATATCTCCCTCCAGGTTCCCACTGCGGATGAGAGGAGCATGACCG 970
 Db 901 TCCGAGCGGATATCTCCCTCCAGGTTCCCACTGCGGATGAGAGGAGCATGACCG 960
 Qy 971 ATCGTGGAGGGGCGCTGCTCGGATGAGAGGCTTGGCAAGTGTGATCTGACATT 1030
 Db 961 ATCGTGGAGGGGCGCTGCTCGGATGAGAGGCTTGGCAAGTGTGATCTGACATT 1020
 Qy 1031 GGGACACCCCACTCTGAGAGGAGCGCTCATGACCGCGAGTGGTGGTGGTGGTGG 1090
 Db 1021 GGGACACCCCACTCTGAGAGGAGCGCTCATGACCGCGAGTGGTGGTGGTGGTGG 1080
 Qy 1091 CACTGCTTCTTCTGAGACCGGGAGAGAGGCTCTGAGAGGCTGGAAGGTTAGCGGG 1150
 Db 1081 CACTGCTTCTTCTGAGACCGGGAGAGAGGCTCTGAGAGGCTGGAAGGTTAGCGGG 1140
 Qy 1151 AGCAACTGACACGAGTGGCTGAGAGCGCTTCATGCGAGATCATCAACAGCAAT 1210
 Db 1141 AGCAACTGACACGAGTGGCTGAGAGCGCTTCATGCGAGATCATCAACAGCAAT 1200
 Qy 1211 TACACCGATGAGAGAGAGAGATATGAGATGCGCCCTCATGAGGCTGTCCAAAGCC 1270
 Db 1201 TACACCGATGAGAGAGAGAGATATGAGATGCGCCCTCATGAGGCTGTCCAAAGCC 1260

Qy 1271 CTGTCCGCTCATCAACCTGCTTGGCTCCCAATGATGAGACAGACTTTAGCTCAAT 1330
 Db 1261 CTGTCCGCTCATCAACCTGCTTGGCTCCCAATGATGAGACAGACTTTAGCTCAAT 1320
 Qy 1331 GAGACTGCTGGATCAAGGCTTTGGCAAGCAGGAGAGAGATGACAGACATCCCC 1390
 Db 1321 GAGACTGCTGGATCAAGGCTTTGGCAAGCAGGAGAGAGATGACAGACATCCCC 1380
 Qy 1391 TTCTCCGGAGGAGGATGATCTGATGAGATCACTTCAAGAAATGATGATGATGAT 1450
 Db 1381 TTCTCCGGAGGAGGATGATCTGATGAGATCACTTCAAGAAATGATGATGATGAT 1440
 Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGAT 1510
 Db 1441 TATGACAGTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGAT 1500
 Qy 1511 TCTGCGAGGAGAGAGAGGCGGCTCTGCTGCTGATGAGAGAGAGAGAGAGAG 1570
 Db 1501 TCTGCGAGGAGAGAGAGGCGGCTCTGCTGCTGATGAGAGAGAGAGAGAGAG 1560
 Qy 1571 GCAGTGTCAACAGCTGGGAGAGAGGCTGCTGCTGCTGAGAGAGAGAGAGAGAG 1630
 Db 1561 GCAGTGTCAACAGCTGGGAGAGAGGCTGCTGCTGCTGAGAGAGAGAGAGAGAG 1620
 Qy 1631 AAAGTGAAGAGAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGC 1675
 Db 1621 AAAGTGAAGAGAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGC 1665

RESULT 13
 BC018715
 LOCUS
 DEFINITION Homo sapiens transmembrane protease, serine 13, mRNA (cDNA clone IMAGE:3505123).
 VERSION BC018715
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klapper, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepien, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Lott, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McKernan, P.J., McKernan, K.D., Malek, U.A., Gunaratne, P.H., Richards, S., Wozniak, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smolins, D.E., Mammalian Gene Collection Program Team
 Mammalian Gene Collection Program Team
 human and mouse cDNA sequences of more than 15,000 full-length
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3222)

NIH GGC Project
 Direct Submission
 Submitted (07-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Bethesda, MD 20892-2590, USA

LOCUS	AB048797	3324 bp	mRNA	linear	PRI 22-MAR-2001
DEFINITION	Homo sapiens mps mRNA for mosaic serine protease, complete cds.				
ACCESSION	AB048797				
VERSION	AB048797.1	GI:13429971			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (sites)				
AUTHORS	Kim,D.R., Sharmir,S., Inoue,M. and Kido,H.				
TITLE	Cloning and expression of novel mosaic serine proteases with and without a transmembrane domain from human lung				
JOURNAL	Biochim. Biophys. Acta 1518 (1-2), 204-209 (2001)				
PUBMED	11267681				
REFERENCE	2 (bases 1 to 3324)				
AUTHORS	Kim,D.R., Inoue,M. and Kido,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-SEP-2000) Hiroshi Kido, The University of Tokushima, Division of Enzyme Chemistry, Institute for Enzyme Research, Kuramoto-cho, Tokushima 770, Japan (E-mail:kido@ier.tokushima-u.ac.jp, Tel:81-88-633-7424, Fax:81-88-633-7425)				
FEATURES	Location/Qualifiers				
source	1..3324				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/tissue_type="lung"				
	/clone_lib="Marathon Ready cDNA (Clontech)"				
gene	1..3324				
	/gene="mps"				
5'UTR	1..87				
	/gene="mps"				
CDS	88..1701				
	/gene="mps"				
	/note="non-transmembrane type"				
	/codon_start=1				
	/transl_except="(pos:787..789,aa:Arg)				
	/transl_except="(pos:904..906,aa:His)				
	/transl_except="(pos:1498..1500,aa:His)				
	/product="mosaic serine protease"				
	/protein_id="BAB39742.1"				
	/db_xref="GI:13429972"				
	/translation="MERSDHGNSAPRTSPASPAQSPAGTPPGASPAQSPASPAQ SPAGTPPGASPAQSPAGTPPGASPAQSPASPAQSPASPAQSPASPAQSPASPAQ RASSGSSBASASVTTSPTRVIVATPVCAVPIRSSASPAATATRESPTVPMQ GHGIRYKKEORSCPKAVRCGVVDCTKSPDELGVCFDMDKSLKIVSSSHQWLP ICSSNWDVSEKTRQDGFSAHRTTEVAHRDFANSFLLRVNSTIOESLHRSCHPS QRIISLOCSHGLAMTGRIVGALASDKPMQVSLHPGTHICGGTLIDQWVLA AHCFVTRKEVLEGMKVYAGTSNHLQPEASIAEIIINSYTDSEDDYDIALMLSK PLTSAIHPACPLPMHCOTPSINFTCWITGPKTRRFDQNTSPILREVNQVLDLFFKC NDIYVDSYLTFRPMCMAGDLHGGRDSCGDSGGLVCCQNNRMVLAGTISNGTGGQR NKKGVYTKTVLEWLVFYSKMSREVRFRKS"				
	1702..3324				
	/gene="mps"				
	3273..3278				
	/gene="mps"				
ORIGIN	polyA_signal				
	3'UTR				
	Query Match 84.6%; Score 1479; DB 5; Length 3324; Best Local Similarity 92.4%; Pred. No. 0; Matches 1634; Conservative 0; Mismatches 0; Indels 135; Gaps 2;				
QY	1	CTCAGAGACCATGAGAGGACAGCCAGGGAATGATCTCCAGACAGAAACACTTCACG	60		
DB	78	CTCAGAGACCATGAGAGGACAGCCAGGGAATGATCTCCAGACAGAAACACTTCACG	137		
QY	61	TGAGACATCTCCAGCCAGGACATCTCCAGTGGGACACCTCCAGGCGGGCATCTCCAGC	120		
DB	138	TGAGACATCTCCAGCCAGGACATCTCCAGTGGGACACCTCCAGGCGGGCATCTCCAGC	197		
QY	121	CCAGGATCTCCAGCCAGGACATCTCCAGTGGGACACCTCCAGGCGGGCATCTCCAGC	180		
DB	198	CCAGGATCTCCAGCCAGGACATCTCCAGTGGGACACCTCCAGGCGGGCATCTCCAGC	257		
QY	181	CCAGGATCTCCAGCTGGTACACTCCAGGCGGGCATCTCCAGGCGG-----	228		
DB	258	CCAGGATCTCCAGCTGGTACACTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGC	317		
QY	229	-----GGCATCTCCAGCCAGGACATCTCCAGGCGGGCATCTCCAGC	270		
DB	318	CCAGGATCTCCAGCCAGGACATCTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGC	377		
QY	271	TTGGGATCACTTCCAGGCTCTTCATCCGAGGTATCATCCGACAGTCAAGCTTCGCT	330		
DB	378	TTGGGATCACTTCCAGGCTCTTCATCCGAGGTATCATCCGACAGTCAAGCTTCGCT	437		
QY	331	GACAACTCTCCCAACAGAGTGTACTTTGTAGAGCAACCAATGGGGGTGTACCAT	390		
DB	438	GACAACTCTCCCAACAGAGTGTACTTTGTAGAGCAACCAATGGGGGTGTACCAT	497		
QY	391	CCGATCATCTCCGACAGTCAAGCAGGACCAAGGAGGACAGGAGAGCCAGGATAC	450		
DB	498	CCGATCATCTCCGACAGTCAAGCAGGACCAAGGAGGACAGGAGAGAGCCAG---	553		
QY	451	GAGCTTCCCAAGTTCACTGGCGGGAGGACCAAGAGACGTACCGCTCATGGGTGCT	510		
DB	554	-----	553		
QY	511	GCTCTCTCTCATTTGCCCTGTGTGTTTCCGCTCATCTCTCTTCCAGTTTGGAGGCGCA	570		
DB	554	-----TCGAGTTTGGAGGCGCA	572		
QY	571	CACAGGATCAGGTACAGAGGACAGAGGAGAGGTGTCCCAAGCACGCTGTGCTGTGA	630		
DB	573	CACAGGATCAGGTACAGAGGACAGAGGAGAGGTGTCCCAAGCACGCTGTGCTGTGA	632		
QY	631	CGGGGTGTGAGCTGCAAGCTGAAAGTGAAGAGCTGGCTGCTGAGTTTGACTGGA	690		
DB	633	CGGGGTGTGAGCTGCAAGCTGAAAGTGAAGAGCTGGCTGCTGAGTTTGACTGGA	692		
QY	691	CAAGTCTCTGTTAAATCTACTCTGGGTCTCCCATAGTGGCTTCCATCTGTACAG	750		
DB	693	CAAGTCTCTGTTAAATCTACTCTGGGTCTCCCATAGTGGCTTCCATCTGTACAG	752		
QY	751	CACTGGAATGACTCTCACTCAGAGAAAGCTGCGACAGCTGGGTTTCAGAGATGCTCA	810		
DB	753	CACTGGAATGACTCTCACTCAGAGAAAGCTGCGACAGCTGGGTTTCAGAGATGCTCA	812		
QY	811	CCGACCAACGAGGTTGCCACAGGATTTTCCCAACAGCTTCTCAATCTTGAGATCAA	870		
DB	813	CCGACCAACGAGGTTGCCACAGGATTTTCCCAACAGCTTCTCAATCTTGAGATCAA	872		
QY	871	CTCAGCATCAGAGAAAGCTTCCACAGTGTGAATGCGCTTCCAGGCGGTATATCTCCT	930		
DB	873	CTCAGCATCAGAGAAAGCTTCCACAGTGTGAATGCGCTTCCAGGCGGTATATCTCCT	932		
QY	931	CCAGTGTTCCTCAGCTGCGGACCTGAGGCGCATGATCGTGGAGGCGGCTGCG	990		
DB	933	CCAGTGTTCCTCAGCTGCGGACCTGAGGCGCATGATCGTGGAGGCGGCTGCG	992		
QY	991	CTCGATAGCAAGTGGCTTGGCAAGTGAATCTTCCGACCACTTCTGTGG	1050		
DB	1051	CTCGATAGCAAGTGGCTTGGCAAGTGAATCTTCCGACCACTTCTGTGG	1052		
QY	1051	AGGACGCTCATTTGAGCGCCAGTGGGTGTCTCATGCGCGCCCATGCTTTCGGAACCG	1110		
DB	1053	AGGACGCTCATTTGAGCGCCAGTGGGTGTCTCATGCGCGCCCATGCTTTCGGAACCG	1112		
QY	1111	GGAGAGGCTCTGAGAGGCTGGAAGTGTACGCGGAGCAGCAACCTGTGACAGATTGCC	1170		
DB	1113	GGAGAGGCTCTGAGAGGCTGGAAGTGTACGCGGAGCAGCAACCTGTGACAGATTGCC	1172		
QY	1171	TGAGGACGCTCATTTGGCGAGATCATCAACAGCAATTACCGATGAGAGAGCA	1230		


```

Db      1173 TGAGGCGAGCTCCATTCGCGAGATCATCTAAGCAATTACCGATAGAGAGAGCA 1232
Qy      1231 CTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGACCTCTGCGCTACATCCACC 1290
Db      1233 CTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGACCTCTGCGCTACATCCACC 1292
Qy      1291 TGCTTGCTCCCAATGATGAGACGACTTATGCTCAATGAGACCTGCTGATCACAAG 1350
Db      1293 TGCTTGCTCCCAATGATGAGACGACTTATGCTCAATGAGACCTGCTGATCACAAG 1352
Qy      1351 CTTTGCGAAGACCGAGGAGACAGATGACAAGACATCCCTCTCCGCGAGGTGACAGT 1410
Db      1353 CTTTGCGAAGACCGAGGAGACAGATGACAAGACATCCCTCTCCGCGAGGTGACAGT 1412
Qy      1411 CAATCTCATGACTTCAAGAAATGCAATGACTACTTGGTCTATGACAGTTACCTTACCCC 1470
Db      1413 CAATCTCATGACTTCAAGAAATGCAATGACTACTTGGTCTATGACAGTTACCTTACCCC 1472
Qy      1471 AAGATGATGTGTGCTGCGGAGCTTGTGCGGAGAGACCTCCGCGAGGAGACAGCG 1530
Db      1473 AAGATGATGTGTGCTGCGGAGCTTGTGCGGAGAGACCTCCGCGAGGAGACAGCG 1532
Qy      1531 GGGGCTCTTGTCTGTGAGCAGAACAAACCGCTGTGCTCTGCGAGGTGACAGCTGCGG 1590
Db      1533 GGGGCTCTTGTCTGTGAGCAGAACAAACCGCTGTGCTCTGCGAGGTGACAGCTGCGG 1592
Qy      1591 CACAGGCTGTGCGCAGAGAAACAAACCTGTGTGTACACCAAGTGAACAGAGTTCTTCC 1650
Db      1593 CACAGGCTGTGCGCAGAGAAACAAACCTGTGTGTACACCAAGTGAACAGAGTTCTTCC 1652
Qy      1651 CTGATTTTACAGCAAGATGAGAGCGAGGTGCGATTCAGAAATCTTAACTGAGTGGCTT 1710
Db      1653 CTGATTTTACAGCAAGATGAGAGCGAGGTGCGATTCAGAAATCTTAACTGAGTGGCTT 1712
Qy      1711 GCTGCTCTGACAGACACCGGCTGCTGTGA 1739
Db      1713 GCTGCTCTGACAGACACCGGCTGCTGTGA 1741

```

```

RESULT 15
AX360089 1614 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 45 from Patent W0200860.
ACCESSION AX360089
VERSION AX360089.1 GI:18675715
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

```

REFERENCE
AUTHORS Plowman, G., Whyte, D., Sudarsanam, S., Manning, G., Caenepeel, S. and
Charyczak, G.
TITLE Novel proteases
JOURNAL Patent: WO 020860-A 45 03-JAN-2002;
Sugen, Inc. (US)
FEATURES
source 1. 1614
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

ORIGIN

```

Query Match 81.8%; Score 1429; DB 2; Length 1614;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 0; Indels 135; Gaps 2;
Qy 11 ATGAGAGGAGGACAGCCAGGGAATGCACTTCACAGCAAGAACACCTTCAGCTGAGCATCT 70
Db 1 ATGAGAGGAGGACAGCCAGGGAATGCACTTCACAGCAAGAACACCTTCAGCTGAGCATCT 60

```

```

Qy      71 CCAGCCAGGACATCTCCAGCTGGGACACATCTCAGAGCCGGGCAATCTCCAGCCAGGACATCT 130
Db      61 CCAGCCAGGACATCTCCAGCTGGGACACATCTCAGAGCCGGGCAATCTCCAGCCAGGACATCT 120
Qy      131 CCAGCCAGGACATCTCCAGCTGGGACACATCTCAGAGCCGGGCAATCTCCAGCCAGGACATCT 190
Db      121 CCAGCCAGGACATCTCCAGCTGGGACACATCTCAGAGCCGGGCAATCTCCAGCCAGGACATCT 180
Qy      191 CCAGCTGTACACCTCCAGGCGGGGCAATCTCCAGGCGG----- 228
Db      181 CCAGCTGTACACCTCCAGGCGGGGCAATCTCCAGGCGGCGGCAATCTCCAGCCAGGACATCT 240
Qy      229 -----GGCATCTCAGGCCAGGACATCTCAGGCCGGGCAATCTCCAGCCAGGACATCT 280
Db      241 CCAGCCAGGACATCTCCAGGCCAGGACATCTCAGGCCGGGCAATCTCCAGCCAGGACATCT 300
Qy      281 CTTTCCAGGCTCTTCATCCGCGAGGTATCATCCGCGAGGTGACGCTTGGTGAACAATCTCC 340
Db      301 CTTTCCAGGCTCTTCATCCGCGAGGTATCATCCGCGAGGTGACGCTTGGTGAACAATCTCC 360
Qy      341 CCAGCCAGGACATCTCCAGGCGGACACACAGTGGGGGCTGTACCAATCCGATCATCT 400
Db      361 CCAGCCAGGACATCTCCAGGCGGACACACAGTGGGGGCTGTACCAATCCGATCATCT 420
Qy      401 CTTGCGAGGTGACGACAGGACACAGGCGGACACAGGAGAGCCAGGTACGAGCTGCGCC 460
Db      421 CTTGCGAGGTGACGACAGGACACAGGCGGACACAGGAGAGCCAG----- 466
Qy      461 AAGTTCACTTGCGGAGGAGGACAGAGACAGCTACCGGTATCGGAGTGCTGCTCTCTC 520
Db      467 ----- 466
Qy      521 ATTGCCCTGTGTTTGGCTCTCATCATCTCTTCCAGTTCTTGCGAGGCGCACACAGGATC 580
Db      467 -----TCAGTTCTTGCGAGGCGCACACAGGATC 495
Qy      581 AGGTACAGGAGGACAGAGAGAGAGCTGTCCAGAGCAGCTGTTGCTGTGACGCGGCTGAT 640
Db      496 AGGTACAGGAGGACAGAGAGAGAGCTGTCCAGAGCAGCTGTTGCTGTGACGCGGCTGAT 555
Qy      641 GACTGCAAGCTGAAGAGATGACAGAGCTGGGCTGCGTGAAGTTTGACTGGGACAACTCTTG 700
Db      556 GACTGCAAGCTGAAGAGATGACAGAGCTGGGCTGCGTGAAGTTTGACTGGGACAACTCTTG 615
Qy      701 CTTAAATCTACTTGTGGCTCTCCATCAGTGGCTTTCCCATCTGTACAGGAACTGGAAAT 760
Db      616 CTTAAATCTACTTGTGGCTCTCCATCAGTGGCTTTCCCATCTGTGTACAGGAACTGGAAAT 675
Qy      761 GACTCTTACTCAGAGAGACCTGCGACAGCTGGGTTTGCAGAGTGTCAACCGGACAAC 820
Db      676 GACTCTTACTCAGAGAGACCTGCGACAGCTGGGTTTGCAGAGTGTCAACCGGACAAC 735
Qy      821 GAGGTTGCCACAGAGGATTTTGCACACAGCTTCTCAATCTTGAGATACAACTCCACATC 880
Db      736 GAGGTTGCCACAGAGGATTTTGCACACAGCTTCTCAATCTTGAGATACAACTCCACATC 795
Qy      881 CAGGAAGCTCCACAGAGTCTGAATGCCCTTCCAGGCGGTATATCTCCTTCAATGTTC 940
Db      796 CAGGAAGCTCCACAGAGTCTGAATGCCCTTCCAGGCGGTATATCTCCTTCAATGTTC 855
Qy      941 CACTGCGGACGTGAGGCGCATGACCGGCGGATGCTGAGGAGGCGGCTGAGCTCGGATAGC 1000
Db      856 CACTGCGGACGTGAGGCGCATGACCGGCGGATGCTGAGGAGGCGGCTGAGCTCGGATAGC 915
Qy      1001 AAGTGGCTTTGGCAAGTGAAGTCTGCACTTGGCACAACCAATCTGTGAGGAGCGCTC 1060
Db      916 AAGTGGCTTTGGCAAGTGAAGTCTGCACTTGGCACAACCAATCTGTGAGGAGCGCTC 975
Qy      1061 ATTGAGCCCAAGTGGGTCTCACTGCGCGCCCACTGCTTTCTTGCGTACCCCGGGAAGAGTC 1120
Db      976 ATTGAGCCCAAGTGGGTCTCACTGCGCGCCCACTGCTTTCTTGCGTACCCCGGGAAGAGTC 1035
Qy      1121 CTGAGGAGGTGAGAGGTGTACGCGGAGACAGCAACTGCACTCAAGTTGCTGAGGAGGCC 1180

```

```
Db 1036 CTGGAGGAGCTGGAGGAGGTGTAGCGGGGCAACGAGCAACCTGCAACGATTGCCGTGAGGCGAGCC 1095
QY 1181 TCCATTGCCGAGATCATCATCAACAGCAATTAACCCGATGAGAGAGCACTATGACATC 1240
Db 1096 TCCATTGCCGAGATCATCATCAACAGCAATTAACCCGATGAGAGAGCACTATGACATC 1155
QY 1241 GCCCTCATGCGGCTGTCCAAAGCCCTGACCCCTGTCCGCTCAATCAACCTGTGCTGCTC 1300
Db 1156 GCCCTCATGCGGCTGTCCAAAGCCCTGACCCCTGTCCGCTCAATCAACCCCTGTGCTGCTC 1215
QY 1301 CCCATGATGAGACGACCTTTAGCTTCAATGAGACCTGTGATCACAAGCTTTGGCAAG 1360
Db 1216 CCCATGATGAGACGACCTTTAGCTTCAATGAGACCTGTGATCACAAGCTTTGGCAAG 1275
QY 1361 ACCAGGAGACAGATGACAAGACATCCCTTCTCCGAGAGTGACAGGTCAATCTCATC 1420
Db 1276 ACCAGGAGACAGATGACAAGACATCCCTTCTCCGAGAGTGACAGGTCAATCTCATC 1335
QY 1421 GACTTCAAGAAATGCAATGACTTGTGTATGACAGTTACTTACCCCAAGGATGATG 1480
Db 1336 GACTTCAAGAAATGCAATGACTTGTGTATGACAGTTACTTACCCCAAGGATGATG 1395
QY 1481 TGTGCTGGGGACCTTCTGTGGGGGAGAGACTCTGCGCAGGAGACAGCGGGGGGCTCTT 1540
Db 1396 TGTGCTGGGGACCTTCTGTGGGGGAGAGACTCTGCGCAGGAGACAGCGGGGGGCTCTT 1455
QY 1541 GTCTGTAGAGCAACAACCGCTGTACTGTGACAGGTGTCAACAGCTGGGGCACAAGCTGT 1600
Db 1456 GTCTGTAGAGCAACAACCGCTGTACTGTGACAGGTGTCAACAGCTGGGGCACAAGCTGT 1515
QY 1601 GGCCAGAGAAACAAACCTGTGTGTACACCAAGTGAAGAAAGTTCTTCCCTGGAATTTAC 1660
Db 1516 GGCCAGAGAAACAAACCTGTGTGTACACCAAGTGAAGAAAGTTCTTCCCTGGAATTTAC 1575
QY 1661 AGCAAGATGAGAGCGAGGTGCGATTGAGAAATCCTAA 1699
Db 1576 AGCAAGATGAGAGCGAGGTGCGATTGAGAAATCCTAA 1614
```

Search completed: September 18, 2006, 11:42:09
Job time : 9639 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:45:49 ; Search time 8720 Seconds
(without alignments)
11209.517 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 1748
Sequence: 1 ctccagagaccatgagagaggg.....gctcgtctgaccgcgagaa 1748

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est3:*
- 3: gb_est4:*
- 4: gb_est5:*
- 5: gb_est6:*
- 6: gb_hc:*
- 7: gb_est2:*
- 8: gb_est7:*
- 9: gb_est8:*
- 10: gb_est9:*
- 11: gb_g881:*
- 12: gb_g882:*
- 13: gb_g883:*
- 14: gb_g884:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1221.2	69.9	1472	6	CR613669 full-leng
2	1221.2	69.9	2029	4	CR596543 full-leng
3	969.4	55.5	1039	4	BX356693 BX356693
4	813	46.5	821	4	BX337410 BX337410
5	741	42.4	945	4	BX337409 BX337409
6	711.4	40.7	930	3	BU957076 AGENCOURT
7	709.4	40.6	834	7	BE732381 601569601
8	637.2	36.5	1043	3	BU558176 AGENCOURT
9	624	35.7	915	3	BU860003 AGENCOURT
10	621	35.5	693	9	DN998208 TC106402
11	581.8	33.3	595	4	CB270010 1008917 H
12	552.8	31.6	777	4	BX389632 BX389632
13	535.6	30.6	852	9	DA830024 DA830024
14	532.8	30.5	640	4	BX952562 DKFZP781X
15	493.4	28.2	579	3	BP359643 BP359643
16	486.2	27.8	845	9	DN932137 AGENCOURT
17	477.4	27.3	596	3	BP319980 BP319980
18	471	26.9	878	4	BX337187 BX337187
19	470.4	26.9	694	4	CA313327 UT-CF-FNO

20	468	26.8	583	3	BP315554	BP315554
21	468	26.8	590	9	DB290424	DB290424
22	462	26.4	586	9	DA848528	DA848528
23	453	25.9	575	9	DB223460	DB223460
24	452.8	25.9	579	9	DA445009	DA445009
25	449.8	25.7	583	3	BP315772	BP315772
26	440.8	25.2	678	7	BE289529	BE289529
27	427	24.4	748	9	DN948363	DN948363
28	411.6	23.5	614	7	BE286322	BE286322
29	386.4	22.1	679	5	CP899959	CP899959
30	359.2	20.5	377	5	CD633581	CD633581
31	358.8	20.5	622	7	BF119003	BF119003
32	357.6	20.5	378	5	CD633580	CD633580
33	355.4	20.3	364	8	CN335419	CN335419
34	348	19.9	402	10	R78581	R78581
35	335.8	19.2	621	7	BE285038	BE285038
36	298.8	17.1	1106	4	BX356692	BX356692
37	286	16.4	333	4	CA406437	CA406437
38	285	16.3	285	7	AW845106	AW845106
39	276.4	15.8	286	1	A1909842	A1909842
40	266.4	15.2	563	7	BE290088	BE290088
41	265.4	15.2	751	7	BE280394	BE280394
42	254.4	14.6	582	3	BP356834	BP356834
43	240.2	13.7	252	2	BF757612	BF757612
44	240.2	13.7	254	2	BF757611	BF757611
45	240.2	13.7	255	2	BF757610	BF757610

ALIGNMENTS

RESULT 1
CR613669
LOCUS
DEFINITION
full-length cDNA clone CSOD1041Y07 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR613669
VERSION
HTC; CNSUT_CDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1472)
REFERENCE
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifestech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1472)
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
BP 191 91006 EYRY cedex - FRANCE (E-mail: segre@genoscope.cns.fr
Web: www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1041Y07"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 69.9%; Score 1221.2; DB 6; Length 1472;

Best Local Similarity 98.6%; Pred. No. 2e-281;
Matches 1248; Conservative 0; Mismatches 3; Indels 15; Gaps 1;

```

Qy 31 GAATGCACTCTCCAGCAAGAAACACCTTCAGCTGAGGAGCATCTCCAGCCCGGAGCATCTCCAGC 90
Db 1 GAATGCACTCTCCAGCAAGAAACACCTTCAGCTGAGGAGCATCTCCAGCCCGGAGCATCTCCAGC 60
Qy 91 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 150
Db 61 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 120
Qy 151 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCTGATACCTCCAGC 210
Db 121 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCTGATACCTCCAGC 180
Qy 211 CCGGGGCACTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 255
Db 181 CCGGGGCACTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 240
Qy 256 CCGGGGCACTCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 315
Db 241 CCGGGGCACTCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 300
Qy 316 CAGGTCAAGCTCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 375
Db 301 CAGGTCAAGCTCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 360
Qy 376 GGGGGGCTGATCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 435
Db 361 GGGGGGCTGATCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 420
Qy 436 GGAAGACCCAGGATGAGAGCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 495
Db 421 GGAAGACCCAGGATGAGAGCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 480
Qy 496 GCTCATCGGGGCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 555
Db 481 GCTCATCGGGGCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 540
Qy 556 GTTCGGAGGAGGACCAAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 615
Db 541 GTTCGGAGGAGGACCAAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 600
Qy 616 CGCTGTTGCTGTGACCGGGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 675
Db 601 CGCTGTTGCTGTGACCGGGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 660
Qy 676 GAGGTTTGAATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 735
Db 661 GAGGTTTGAATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 720
Qy 736 TCCCATCTGTAGAGAGCACTGGAATGATCTCCAGCTCAGAGAGAGCTCCAGAGAGCTCCAGAGCTCCAGG 795
Db 721 TCCCATCTGTAGAGAGCACTGGAATGATCTCCAGCTCAGAGAGAGCTCCAGAGAGCTCCAGAGCTCCAGG 780
Qy 796 TTTTGAGAGTGTCTACCGGAGCAACCGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 855
Db 781 TTTTGAGAGTGTCTACCGGAGCAACCGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 840
Qy 856 AATCTTGAGATACAACTCCAGCTCAGAGAAAGCTCCAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 915
Db 841 AATCTTGAGATACAACTCCAGCTCAGAGAAAGCTCCAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 900
Qy 916 GCGGATATCTCTCCAGCTCCAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 975
Db 901 GCGGATATCTCTCCAGCTCCAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 960
Qy 976 GGGAGGGGCGCTGAGCTGAGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1035
Db 961 GGGAGGGGCGCTGAGCTGAGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1020
Qy 1036 CACCCACATCTGTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1095

```

```

Db 1021 CACCCACATCTGTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1080
Qy 1096 CTTCTGTGATGAGCCGGGAGAGAGGATCTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1155
Db 1081 CTTCTGTGATGAGCCGGGAGAGAGGATCTGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1140
Qy 1156 CTTGACACAGTGTGAGGAGGATCTCCAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1215
Db 1141 CTTGACACAGTGTGAGGAGGATCTCCAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1200
Qy 1216 CGATGAGAGAGGACATATGATGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1275
Db 1201 CGATGAGAGAGGACATATGATGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1260
Qy 1276 CGCTCA 1281
Db 1261 CGGTGA 1266

```

RESULT 2
CR596543
LOCUS
DEFINITION full-length cDNA clone CS0D1015YU24 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR596543
VERSION CR596543.1 GI:50477350
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2029)

REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqret@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
Location/Qualifiers
source
1..2029
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YU24"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 69.9%; Score 1221.2; DB 6; Length 2029;
Best Local Similarity 98.6%; Pred. No. 2e-281;
Matches 1248; Conservative 0; Mismatches 3; Indels 15; Gaps 1;

```

Qy 31 GAATGCACTCTCCAGCAAGAAACACCTTCAGCTGAGGAGCATCTCCAGCCCGGAGCATCTCCAGC 90
Db 1 GAATGCACTCTCCAGCAAGAAACACCTTCAGCTGAGGAGCATCTCCAGCCCGGAGCATCTCCAGC 60
Qy 91 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 150
Db 61 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGC 120
Qy 151 TGGGACACCTCCAGCCCGGAGCATCTCCAGCCCGGAGCATCTCCAGCTGATACCTCCAGC 210

```

```

Db      121 TGGAGACACCTCCGGGCGGGGATCTCCAGCCAGGCAATCTCAGCTGGTACATCTCCAGG 180
Qy      211 CCGGGGCACTCTCCAGCCG-----GGCATCTCCAGCCAGGCAATCTCCAGC 255
Db      181 CCGGGGCACTCTCCAGCCGCGGGATCTCCAGCCAGGCAATCTCCAGC 240
Qy      256 CCGGGGCACTCTCCAGCTCCGATCACTTCCAGGCTCCATCCGCGAGGATCATCTCCG 315
Db      241 CCGGGGCACTCTCCGCTCTGGGATCACTTCCAGGCTCCATCCGCGAGGATCATCTCCG 300
Qy      316 CAGGTCAAGCTCCGATGACCACTCCCAACCAAGTGTACCTTTAGAGCAACCAAGT 375
Db      301 CAGGTCAAGCTCCGATGACCACTCCCAACCAAGTGTACCTTTAGAGCAACCAAGT 360
Qy      376 GGGGGGCTGTACCTCCGATCATCTCTGCCAGGTCAAGCAACCAAGGCGCCAGC 435
Db      361 GGGGGGCTGTACCTCCGATCATCTCTGCCAGGTCAAGCAACCAAGGCGCCAGC 420
Qy      436 GGAAGGCCGAGGTACAGGCTCCAGGCTCCAGGCTCCAGGCGCGAGGCGGAGGAGTACC 495
Db      421 GGAAGGCCGAGGTACAGGCTCCAGGCTCCAGGCTCCAGGCGCGAGGCGGAGGAGTACC 480
Qy      496 GCTCATCGGGTGGCTGCTCTCTCATTTGCGTGGTTCGCTCATCATCTCTTCCA 555
Db      481 GCTCATCGGGTGGCTGCTCTCTCATTTGCGTGGTTCGCTCATCATCTCTTCCA 540
Qy      556 GTTCTGAGAGGCGCACAGGGATCAGGTACAGAGAGAGAGAGAGAGTCTCCAGCA 615
Db      541 GTTCTGAGAGGCGCACAGGGATCAGGTACAGAGAGAGAGAGAGTCTCCAGCA 600
Qy      616 CGCTGTCCTGATACCGGGGAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 675
Db      601 CGCTGTCCTGATACCGGGGAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660
Qy      676 GAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 735
Db      661 GAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720
Qy      736 TCCCATCTGTAGAGCAACTGGAATGACTCTCACTCAGAGAGAGAGAGAGAGTGGG 795
Db      721 TCCCATCTGTAGAGCAACTGGAATGACTCTCACTCAGAGAGAGAGAGAGTGGG 780
Qy      796 TTTTGAAGTGTCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855
Db      781 TTTTGAAGTGTCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy      856 AATCTTGAGATACCACTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
Db      841 AATCTTGAGATACCACTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy      916 GCGGATATATCTCCCTCCAGTGTCTCCAGTGTCTCCAGTGTCTCCAGTGTCTCCAGT 975
Db      901 GCGGATATATCTCTCTCCAGTGTCTCCAGTGTCTCCAGTGTCTCCAGTGTCTCCAGT 960
Qy      976 GGGAGGGGGGCTGGCTCTGATACCAATGTGCTTTGGAGAGTGTGCACTTCGGGAC 1035
Db      961 GGGAGGGGGGCTGGCTCTGATACCAATGTGCTTTGGAGAGTGTGCACTTCGGGAC 1020
Qy      1036 CACCCACATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1095
Db      1021 CACCCACATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy      1096 CTTCTCTGTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1155
Db      1081 CTTCTCTGTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy      1156 CTTCTCTGTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
Db      1141 CTTCTCTGTGACCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Qy      1216 CGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1275

```

```

Db      1201 CGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy      1276 CGCTCA 1281
Db      1261 CGGTGA 1266

RESULT 3
BX356693
LOCUS
DEFINITION
BX356693 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0D1015YL24 5-PRIME, mRNA sequence.
BX356693
ACCESSION
BX356693
VERSION
BX356693.2 GI:46290274
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1039)
REFERENCE
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30372079.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6066.i
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0D1015DF12QP1&c=6066.i.
FEATURES
source
location/Qualifiers
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YL24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 55.5%; Score 969.4; DB 4; Length 1039;
Best Local Similarity 97.9%; Pred. No. 4.3e-221;
Matches 1002; Conservative 10; Mismatches 7; Indels 4; Gaps 3;
Qy      31 GAATGATCTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 90
Db      1 GAATGATCTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy      91 TGGAGACCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 150
Db      61 TGGAGACCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 120
Qy      151 TGGAGACCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 210
Db      121 TGGAGACCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 180
Qy      211 CCGGGGCACTCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 270
Db      181 CCGGGGCACTCTCCAGGCGGGGATCTCCAGGCGAGGATCTCCAGGCGAGGATCTCCAGC 240
Qy      271 TTTGAGATCACTTCCAGGCTCTCATCCGAGGATCAATCCGCGAGGATCAATCCGCTCGT 330

```

Db 241 TCTGGCATCACTTTCCAGTCTCCATCCGAGGATCATCCGCGAGTCAAGCTCGGT 300
 Oy 331 GACAACCTCCCAACCAAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 390
 Db 301 GACAACCTCCCAACCAAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 360
 Oy 391 CCGATCATCTCTGCGAGTCAAGCAACCAAGGCGCAACAGGAGAGCCAGGTAC 450
 Db 361 CCGATCATCTCTGCGAGTCAAGCAACCAAGGCGCAACAGGAGAGCCAGGTAC 420
 Oy 451 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGCTACCGCTCATTCGGGTGCT 510
 Db 421 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGCTACCGCTCATTCGGGTGCT 480
 Oy 511 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGGAGGCGCA 570
 Db 481 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGGAGGCGCA 540
 Oy 571 CACAGGATCAGGTACAAAGAGCAGAGGAGAGGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 630
 Db 541 CACAGGATCAGGTACAAAGAGCAGAGGAGAGGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 600
 Oy 631 CCGGGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCTGAGAGTTGACTGGGA 690
 Db 601 CCGGGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCTGAGAGTTGACTGGGA 660
 Oy 691 CAAGCTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTGAGAG 750
 Db 661 CAAGCTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTGAGAG 720
 Oy 751 CAAGCTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTGAGAG 810
 Db 721 CAAGCTCTGCTTAAATCTACTCTGGGTCTCCCATCAGTGGCTTCCATCTGTGAGAG 780
 Oy 811 CCGGACCAACGAGGTTGGCCACAGGATTTTGGCAACAGCTTCTCAATCTTGAGATACAA 870
 Db 781 CCGGACCAACGAGGTTGGCCACAGGATTTTGGCAACAGCTTCTCAATCTTGAGATACAA 840
 Oy 871 CTCACCATCCAGGAAAGCTTCCACAGGCTGAAATGACCTTCCAGCGGTATATCTCCCT 930
 Db 841 CTCACCATCCAGGAAAGCTTCCACAGGCTGAAATGACCTTCCAGCGGTATATCTCCCT 900
 Oy 931 CCAAGTGTCCCACTGCGGACTGAGGCGCATGACCGGCGGATCTGTGGAGGGCGCTGGC 990
 Db 901 CCAAGTGTCCCACTGCGGACTGAGGCGCATGACCGGCGGATCTGTGGAGGGCGCTGGC 959
 Oy 991 CTCGGATGCAAGTGGCCTTGGCAAGTGAATCTGCACTTCCGACCAACCATCTGTGG 1050
 Db 960 CTCGGATGCAAGTGGCCTTGGCAAGTGAATCTGCACTTCCGACCAACCATCTGTGG 1016
 Oy 1051 AGG 1053
 Db 1017 GGG 1019

RESULT 4
 BX337410 821 bp mRNA linear EST 07-APR-2004
 LOCUS BX337410 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 DEFINITION clone CS0D1041Y07 5-PRIME, mRNA sequence.
 ACCESSION BX337410
 VERSION BX337410.2 GI:46272986
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS Li, M.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 2, 2003 this sequence version replaced gi:30339607.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 6066.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?cs=CS0D1041AG04QPlac=6066.r.
 Location/Qualifiers

FEATURES
 source
 1..821
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1041Y07"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 46.5%; Score 813; DB 4; Length 821;
 Best Local Similarity 98.3%; Pred. No. 1,38-183;
 Matches 807; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

Db 31 GAATGCACTTCCAGCAAGAACACCTTCAAGCTGAGCATCTCCAGCCAGGATCTCCAGC 90
 Oy 1 GAATGCACTTCCAGCAAGAACACCTTCAAGCTGAGCATCTCCAGCCAGGATCTCCAGC 60
 Db 91 TGGGACACCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCCAGGATCTCCAGC 150
 Oy 61 TGGGACACCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCCAGGATCTCCAGC 120
 Db 151 TGGGACACCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCTGTACCTCCAGG 210
 Oy 121 TGGGACACCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCTGTACCTCCAGG 180
 Db 211 CCGGGCATCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCCGGGATCTCCGGC 270
 Oy 181 CCGGGCATCTCCAGGCGGGGATCTCCAGCCAGGCAATCTCCAGCCGGGATCTCCGGC 240
 Db 271 TCTGGCATCACTTCCAGGCTCCATCCGAGGATCATCTCCGCGAGGCTGGCT 330
 Oy 241 TCTGGCATCACTTCCAGGCTCCATCCGAGGATCATCTCCGCGAGGCTGGCT 300
 Db 331 GACAACCTCCCAACCAAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 390
 Oy 301 GACAACCTCCCAACCAAGAGTGTACTTGTAGAGCAACACAGTGGGGGCTGTACCAT 360
 Db 391 CCGATCATCTCTGCGAGTCAAGCAACCAAGGCGCAACAGGAGAGCCAGGTAC 450
 Oy 361 CCGATCATCTCTGCGAGTCAAGCAACCAAGGCGCAACAGGAGAGCCAGGTAC 420
 Db 451 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGCTACCGCTCATTCGGGTGCT 510
 Oy 421 GAGCTCGCCCAAGTTCACCTGGCGGAGAGGCGCAAGAGCTACCGCTCATTCGGGTGCT 480
 Db 511 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGGAGGCGCA 570
 Oy 481 GCTCCTCTCATTTGCGCTGGTGTGTTGCTCATCTCTTCCAGTCTTGGAGGCGCA 540
 Db 571 CACAGGATCAGGTACAAAGAGCAGAGGAGAGGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 630
 Oy 541 CACAGGATCAGGTACAAAGAGCAGAGGAGAGGAGTGTCCCAAGCAAGCTGTTCGCTGTGA 600
 Db 631 CCGGGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCTGAGAGTTGACTGGGA 690
 Oy 601 CCGGGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCTGAGAGTTGACTGGGA 660

QY 691 CAAGTCTGTGCTTAAATCTACTGGGTCCTCCATCAGTGGCTTCCCATCTGTGAG 750
DB 661 CAAGTCTGTGCTTAAATCTACTGGGTCCTCCATCAGTGGCTTCCCATCTGTGAG 720
QY 751 CACTGAGATGACTCTCTACTGAGAGAGACCTGCGAGAGCTGGTTTGGAGAGTCTCA 810
DB 721 CACTGAGATGACTCTCTACTGAGAGAGACCTGCGAGAGCTGGTTTGGAGAGTCTCA 780
QY 811 CCGAGAACCGAGGTTGCCACAGGAGATTTTGGCCACAGCT 851
DB 781 CCGAGAACCGAGGTTGCCACAGGAGATTTTGGCCACAGCT 821

RESULT 5
BX337409/c 945 bp mRNA linear EST 07-APR-2004
LOCUS BX337409 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CS0D1041YM07 3-PRIME, mRNA sequence.
BX337409
ACCESSION BX337409.2 GI:46271157
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30337605.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6066.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0D1041AG04NP1&c=6066.r.
FEATURES
source
1..945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1041YM07"
/tissue="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 42.4%; Score 741; DB 4; Length 945;
Best Local Similarity 96.8%; Pred. No. 2.4e-166;
Matches 758; Conservative 16; Mismatches 7; Indels 2; Gaps 2;

QY 679 GTTGAAGTGGACAAGTCTCTGCTTAAATCTAAGTGGTCTCCATCAGTGGCTTCC 738
DB 765 GTTGAAGTGGACAAGTCTCTGCTTAAATCTAAGTGGTCTCCATCAGTGGCTTCC 706
QY 739 CATCTGAGCAAGCACTGGAATGAATCTCTACTGAGAGAGACCTGCCAGAGTGGTTT 798
DB 705 CATCTGAGCAAGCACTGGAATGAATCTCTACTGAGAGAGACCTGCCAGAGTGGTTT 646
QY 799 CGAAGTGTCTACCGGACACACCGAGTTTCCCAAGAGATTTTGGCCACACTTCTCAT 858
DB 645 CGAAGTGTCTACCGGACACACCGAGTTTCCCAAGAGATTTTGGCCACACTTCTCAT 586
QY 859 CTTGAGATACAACTCCACATCCAGAAAGGCTCCACAGTCTGAATSCCTTCCACAG 918
DB 585 CTTGAGATACAACTCCACATCCAGAAAGGCTCCACAGTCTGAATSCCTTCCACAG 526
QY 919 GTATATCTCCCTCCAGATGTTTCCCACTGAGAGCTGAGAGGCGGATGCTGGG 978
DB 525 GTATATCTCCCTCCAGATGTTTCCCACTGAGAGCTGAGAGGCGGATGCTGGG 466
QY 979 AGGGCGCTGCTCTCGATACAGAGTGGCTTGGCAAGTGAAGTCTGCACTTCCGACAC 1038
DB 465 AGGGCGCTGCTCTCGATACAGAGTGGCTTGGCAAGTGAAGTCTGCACTTCCGACAC 406
QY 1039 CCACATCTGTGAGAGGACGCTCATTTGACGCGGAGTGGTCTCACTGCGCCACTGCTT 1098
DB 405 CCACATCTGTGAGAGGACGCTCATTTGACGCGGAGTGGTCTCACTGCGCCACTGCTT 346
QY 1099 CTTGCTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGGACACAGCACT 1158
DB 345 CTTGCTGACCCCGGAGAGAGTCTCTGAGAGGCTGGAAGTGTACGCGGACACAGCACT 286
QY 1159 GCACCAAGTGTCTGAGAGGACGCTCATTTGCGGAGATCATCAACAGCAATTACACCGA 1218
DB 285 GMAWAGTGTGATAGAGGACGCTCAATTTCCGAGATCATCAACAGCAATTACACCGA 226
QY 1219 TGAGAGAGAGACATATGACATGCGCCTCATGCGGCTGCAAGCCCGGACCTGTCCG 1278
DB 225 TGAGAGAGAGACATATGACATGCGCCTCATGCGGCTGCAAGCCCGGACCTGTCCG 166
QY 1279 TCA 1281
DB 165 TGA 163

RESULT 6
BU957076 930 bp mRNA linear EST 21-OCT-2002
LOCUS AGNCOURT_10615692 NIH_MGC_107 Homo sapiens CDNA clone
DEFINITION IMAGE:6731027 5', mRNA sequence.
BU957076
ACCESSION BU957076.1 GI:24186636
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 930)
AUTHORS NIH-MGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
Email: cgabbs-remail.nih.gov
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNCM3056 row: m column: 10

FEATURES High quality sequence stop: 653.

Location/Qualifiers

source

1..930

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6731027"

/issue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC 107"

/note="Organ: b19asc; Vector: pOT7; Site_1: EcoRI;

Site_2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 40.7%; Score 711.4; DB 3; Length 930;

Best Local Similarity 97.6%; Pred. No. 3e-159;

Matches 743; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

236 CCAGCCCAAGGATCTCCAGCCCGGCGATCTCCGCTCTGAGCATCTTCCAGGCTCTCA 295
Db 1 CCAGCCCAAGGATCTCCAGCCCGGCGATCTCCGCTCTGAGCATCTTCCAGGCTCTCA 60
Qy 296 TCCGCGAGGTATATATCCGCGAGGTACAGCTCTGCGTGAACCTCCCAACCAAGGTGAC 355
Db 61 TCCGCGAGGTATATATCCGCGAGGTACAGCTCTGCGTGAACCTCCCAACCAAGGTGAC 120
Qy 356 CTTGTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGCGAGGTACCA 415
Db 121 CTTGTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGCGAGGTACCA 180
Qy 416 CCAGCAACCAAGGCGCAACCAAGGAGAGCCCAAGTACAGAGCTTCCCAAGTCACTGCGAG 475
Db 181 CCAGCAACCAAGGCGCAACCAAGGAGAGCCCAAGTACAGAGCTTCCCAAGTCACTGCGAG 240
Qy 476 GAGGCGCGAAGCAAGTACCGCTCATCGGGGCGGTGCTCTCTCATTTGCCCTGTGTT 535
Db 241 GAGGCGCGAAGCAAGTACCGCTCATCGGGGCGGTGCTCTCTCATTTGCCCTGTGTT 300
Qy 536 TCGCTCATCATCTCTCTCTCAAGTTCTGCGAGGCGCAACAGAGGATCAGGTACAGAGGAG 595
Db 301 TCGCTCATCATCTCTCTCTCAAGTTCTGCGAGGCGCAACAGAGGATCAGGTACAGAGGAG 360
Qy 596 AGGAGAGCTGTCCCAAGCAGCTGTGCTGTGAAGGGGTGTGTGACTGCAAGCTGAAG 655
Db 361 AGGAGAGCTGTCCCAAGCAGCTGTGCTGTGAAGGGGTGTGTGACTGCAAGCTGAAG 420
Qy 656 AGTACAGAGCTGTGCGTGTGAGGTTTGAAGTGAAGTCTCTGCTTAAATCTACTCT 715
Db 421 AGTACAGAGCTGTGCGTGTGAGGTTTGAAGTGAAGTCTCTGCTTAAATCTACTCT 480
Qy 716 GGGTCTCTCCATCAGTGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 775
Db 481 GGGTCTCTCCATCAGTGGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 540
Qy 776 AAGACTCTCCAGCAGCTGGGTTTTCAGAGTGTCTCAACCGGACACCGAGGTTGCCACAG 835
Db 541 AAGACTCTCCAGCAGCTGGGTTTTCAGAGTGTCTCAACCGGACACCGAGGTTGCCACAG 600
Qy 836 GATTTTGGCAACAGTTTCTCAATCTTGAAGATACAACTCCACATCCAGAGAAAGCTCTCAC 895
Db 601 GATTTTGGCAACAGTTTCTCAATCTTGAAGATACAACTCCACATCCAGAGAAAGCTCTCAC 660
Qy 896 AGGTCTGAATGCCCTTCCAGAGGATATCT-CCCTCCAGTGTTCCTCACT-GCGGACTGA 953
Db 661 AGGTCTGAATGCCCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCACTGCGGAGACTGA 720
Qy 954 GGGCCATGACCGGGCGATCTGTGGAGGGGCGGTGGCTCG 994

Db 721 GGGCCATGACCGGGCGATCTGTGGAGGGGCGGTGGCTCG 761

RESULT 7 BE732381 834 bp mRNA linear EST 15-SEP-2000

LOCUS BE732381 601569601.F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844129 5',

DEFINITION mRNA sequence.

ACCESSION BE732381 GI:10146373

VERSION BE732381.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo;

1 (bases 1 to 834)

REFERENCE NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: ggaaps-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov

Plate: LNCM540 Row: D Column: 02

High quality sequence stop: 817.

Location/Qualifiers

1..834

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3844129"

/issue_type="choriocarcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_id="NIH_MGC 21"

/note="Organ: placenta; Vector: pOT7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 40.6%; Score 709.4; DB 7; Length 834;

Best Local Similarity 95.4%; Pred. No. 8.9e-159;

Matches 795; Conservative 0; Mismatches 31; Indels 7; Gaps 6;

664 GCTGGGCTGCGTGAAGTTTGAAGTGAAGTCTCTGCTTAAATCTACTTGGGTCTC 723
Db 1 GCTGGGCTGCGTGAAGTTTGAAGTGAAGTCTCTGCTTAAATCTACTTGGGTCTC 60
Qy 724 CCATCAGTGTGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAGAGCTG 783
Db 61 CCATCAGTGTGCTTCCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAGAGCTG 120
Qy 784 CCAGCAGCTGGGTTTTCAGAGTGTCTCAACCGGACCAACGAGGTTGCCACAGGATTTTGC 843
Db 121 CCAGCAGCTGGGTTTTCAGAGTGTCTCAACCGGACCAACGAGGTTGCCACAGGATTTTGC 180
Qy 844 CAACAGCTTCTCAATCTTGAAGATACAACTCCACATCCAGAGAAAGCTCTCACAGGTTCTGA 903
Db 181 CAACAGCTTCTCAATCTTGAAGATACAACTCCACATCCAGAGAAAGCTCTCACAGGTTCTGA 240
Qy 904 ATGCGCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCACTGCGGAGACTGAGAGGCGCATGAC 963
Db 241 ATGCGCTTCCAGAGGATATCTCCCTCCAGTGTTCCTCACTGCGGAGACTGAGAGGCGCATGAC 300

QY 964 CCGGCGATCGTGGAGGGGCGCTGGCTCGGATGCAAGTGGCTTGGCAAGTAGCT 1023
DB 301 CCGGCGATCGTGGAGGGGCGCTGGCTCGGATGCAAGTGGCTTGGCAAGTAGCT 360
QY 1024 GCACCTTGGGCAACCCACATCTGTGAGGAGCGCTCAATTGACGCCAGTGGTGTCTAC 1083
DB 361 GCACCTTGGGCAACCCACATCTGTGAGGAGCGCTCAATTGACGCCAGTGGTGTCTAC 420
QY 1084 TCCGCGCCATCTGCTTCTTGTGACCCGGGAGAGAGTCTGTGAGGAGTGTATGC 1143
DB 421 TCCGCGCCATCTGCTTCTTGTGACCCGGGAGAGAGTCTGTGAGGAGTGTATGC 480
QY 1144 GGGGACGAGAACCTGCACGAGTGGCTGTGAGGAGCGCTTCCATTGGCGAGATCATCA 1203
DB 481 GGGGACGAGAACCTGCACGAGTGGCTGTGAGGAGCGCTTCCATTGGCGAGATCATCA 538
QY 1204 CAGCAATTACACCGATGAGAGAGAGCACTA-TGACATGAGCGCTCATGCGGCTGTCCAGC 1262
DB 539 CAGCAATTACACCGATGAGAGAGAGCACTAATTGACATGCGCTCATGCGGCTGTCCAG 598
QY 1263 CCGTGAACCTGTCCGCTCACTTCCACCTCTGCTTCCCT-CCCCATGACATGACAGCTTT 1321
DB 599 CCGTGAACCTGTCCGCTCACTTCCACCTCTGCTTCCCTCCCATGACATGACAGAGCTTT 658
QY 1322 AGCCTCATGAGACCTGCTGATGACAGGCTTGGCAAGACGAGGAGACAGATGACAG 1381
DB 659 AGCCTCATGAGACCTGCTGATGACAGGCTTGGCAAGACGAGGAGACAGATGACAG 718
QY 1382 ACATCCCCCTTCTCTCGGAGAGTGCAGGTCACTCATGCACTTCAGAAATGCAATGAC 1441
DB 719 ACATCCCCCTTCTCTCGGAGAGTGCAGGTCACTCATGCACTTCAGAAATGCAATGAC 777
QY 1442 TACTTGTCTATGACAGTTACCTTACCCCAAGATGATGTGTGGGAGCT 1494
DB 778 TAACTGTCTATGACAGT--ACCTTACCCCAAGATGATGTGTGGGAGCT 828

RESULT 8 1043 bp mRNA linear EST 16-SEP-2002
LOCUS BUS58176
DEFINITION AGENCOURT 10212529 NIH MGC 107 Homo sapiens cDNA clone
IMAGE:6586431 5', mRNA sequence.
BUS58176
ACCESSION BUS58176.1 GI:22908472
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1043)
NIH-MGC htp://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/MLNU at:
http://image.llnl.gov
Plate: LICM279 row: p column: 15
High quality sequence stop: 489.
Location/Qualifiers
1..1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6586431"
/issue_type="adenocarcinoma, cell line"

FEATURES
source

ORIGIN
Query Match 36.5%; Score 637.2; DB 3; Length 1043;
Best Local Similarity 98.6%; Pred. No. 1.9e-141;
Matches 653; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 107"
/note="Organ: breast; Vector: pORF7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT printing.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

QY 236 CCAGCCGAGCATCTCAGCCCGGCACTTCGCGCTGTGGCATCACTTCCAGTCTCTCA 295
DB 1 CCAGCCGAGCATCTCAGCCCGGCACTTCGCGCTGTGGCATCACTTCCAGTCTCTCA 60
QY 296 TCCGCGAGTGCATATCGCGCAGGTCAGCGCTCGGTGCAACCTCCCAACGAGGTAC 355
DB 61 TCCGCGAGTGCATATCGCGCAGGTCAGCGCTCGGTGCAACCTCCCAACGAGGTAC 120
QY 356 CTGTGAGCAACACAGTGGGAGCTGTACCCATCCGATCATCTCTGCGAGTCAACA 415
DB 121 CTGTGAGCAACACAGTGGGAGCTGTACCCATCCGATCATCTCTGCGAGTCAACA 180
QY 416 CCAGCAACGAGGCGCACGAGGAGAGCCAGATACAGAGCTGCGCAAGTCACTGCGGG 475
DB 181 CCAGCAACGAGGCGCACGAGGAGAGCCAGATACAGAGCTGCGCAAGTCACTGCGGG 240
QY 476 GAGGCGCAAGAGCAGTACCGCTCATGCGGAGCTGCTCCCTCCATGAGCCCTGGTGT 535
DB 241 GAGGCGCAAGAGCAGTACCGCTCATGCGGAGCTGCTCCCTCCATGAGCCCTGGTGT 300
QY 536 TCGCTCATCATCTCTTCCAGTTCTGAGGAGCCACAGAGGATCAGTACAGAGAGAG 595
DB 301 TCGCTCATCATCTCTTCCAGTTCTGAGGAGCCACAGAGGATCAGTACAGAGAGAG 360
QY 596 AGGAGAGAGCTGTCCAGACAGCGCTTTCGCTGTGACGCGGAGTGTGAGCTGCAAGT 655
DB 361 AGGAGAGAGCTGTCCAGACAGCGCTTTCGCTGTGAGCGGAGTGTGAGCTGCAAGT 420
QY 656 AGTACAGAGCTGGGCTCGGTGAGGTTGACTGCGGACAGTCTCTGCTTAAATCTACTCT 715
DB 421 AGTACAGAGCTGGGCTCGGTGAGGTTGACTGCGGACAGTCTCTGCTTAAATCTACTCT 480
QY 716 GGGTCTCCCATCATGAGGCTTCCATCTGTGAGCAAGTGAATGATCTCTACTCAGAG 775
DB 481 GGGTCTCCCATCATGAGGCTTCCATCTGTGAGCAAGTGAATGATCTCTACTCAGAG 540
QY 776 AAGACTGCGCAGCAGCTGGGTTTGAAGTGTGACCGGACCAACCGAGTTGCCCAAG 835
DB 541 AAGACTGCGCAGCAGCTGGGTTTGAAGTGTGACCGGACCAACCGAGTTGCCCAAG 600
QY 836 GATTTGCAACAGCTTCTCA-TCTTGAGATATACTCCACCATCCAGGAAACCTCCA 894
DB 601 GATTTGCAACAGCTTCTCAATTCTTGAGATACCACTCCACCATCCAGGAAAGCTTCC 660
QY 895 CA 896
DB 661 CA 662

RESULT 9 915 bp mRNA linear EST 16-OCT-2002
LOCUS BUS60003
DEFINITION AGENCOURT 10435147 NIH MGC 107 Homo sapiens cDNA clone
IMAGE:6651006 5', mRNA sequence.
BUS60003
ACCESSION BUS60003.1 GI:24044995
VERSION
KEYWORDS
EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 915)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgshpb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LNCM2897 row: c column: 06
High quality sequence atop: 460.
Location/Qualifiers
1..915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:665106"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 107"
/note="Organ: breast; Vector: pOT57; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACACAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

ORIGIN
Query Match 35.7%; Score 624; DB 3; Length 915;
Best Local Similarity 98.2%; Pred. No. 2.7e-138;
Matches 641; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

236 CCAGCCGAGGATCTCCAGCCCGGAGATCTCCGCTCTGGATCATCTTCAGAGTCTCA 295
1 CCAGCCGAGGATCTCCAGCCCGGAGATCTCCGCTCTGGATCATCTTCAGAGTCTCA 60
236 TCCGCGAGGTATCTCCGCGAGGTCTGAGTGAACAACCTCCCAACGAGGTATAC 355
61 TCCGCGAGGTATCTCCGCGAGGTCTGAGTGAACAACCTCCCAACGAGGTATAC 120
356 CTTGTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTCCAGGTACAGA 415
121 CTTGTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTCCAGGTACAGA 180
416 CCAGCAACGAGGCCACGAGGAGAGCCAGGTACAGCTTGCACCAAGTCACTGGCGG 475
181 CCAGCAACGAGGCCACGAGGAGAGCCAGGTACAGCTTGCACCAAGTCACTGGCGG 240
476 GAGGCGCAGAGCAGCTACCGCTATCGGGTGGCTCTCCCTATGGCCCTGGTGGTT 535
241 GAGGCGCAGAGCAGCTACCGCTATCGGGTGGCTCTCCCTATGGCCCTGGTGGTT 300
536 TCGCTCATCATCTCTTCCAGTTCTGGAGGGCCACAGAGGATCAGGTACAGAGAGAG 595
301 TCGCTCATCATCTCTTCCAGTTCTGGAGGGCCACAGAGGATCAGGTACAGAGAGAG 360
596 AGGAGAGACTGTCCCAACAGCTGTTCGCTGTGAACGAGGTGTGAGTCAAGCTGAAG 655
361 AGGAGAGACTGTCCCAACAGCTGTTCGCTGTGAACGAGGTGTGAGTCAAGCTGAAG 420
656 AGTACGAGCTGGGTGGGTGAGTTTGAAGTGGAGCAAGTCTCTGCTTAAATCTACTCT 715

Db 421 AGTACGAGCTGGGCTGGTGAGGTGTGACTGGAGACAAGTCTGCTTAAATCTACTCT 480
Oy 716 GGGTCTCTCCATCATGAGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 775
Db 481 GGGTCTCTCCATCATGAGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTCAGAG 540
Oy 776 AAGACTTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACACCGAGTTGCCACA-G 834
Db 541 AAGACTTCCAGCAGCTGGGTTTGAAGTGTCTACCGGACACCGAGTTGCCACAAG 600
Oy 835 GGATTTTCCCAACGCTTCTCATCTTGAGTACACTCCACCATCCAGGAAA 887
Db 601 GGATTTTCCCAACGCTTCTCATCTTGAGTACACTCCACCATCCAGGAAA 653

RESULT 10
DN998208
LOCUS
DEFINITION
DN998208 693 bp mRNA linear EST 17-MAY-2005
TC106402 Human breast cancer tissue, large insert, pCMV expression
library Homo sapiens cDNA clone TC106402 5' similar to Homo sapiens
transmembrane protease, serine 13 (TMPSR13), mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 693)
Birkett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,
Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,
Zhang, X., Jay, G., and He, W.
High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
unpublished (2005)
CONTACT: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: CDNA@origene.com
This Bst submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
<http://www.origene.com>
Seq primer: pCMV6 Spine forward vector primer, Origene
Technologies Inc.
FEATURES
SOURCE
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC106402"
/tissue_type="Breast Cancer"
/clone_id="Human Breast Cancer"
expression library"
/note="Organ: Mammary gland (cancer tissue); Vector:
pCMV6-Xis; Site 1: EcoRI; Site 2: XhoI/SalI compatible and
ligatio; oligo-dT primed reverse transcription optimized
for large and GC rich mRNA transcripts; cDNA size
selection, optimized ligation for large inserts into
mamalian expression vector; random clones selected for
end sequence verification of full-length genes"

ORIGIN
Query Match 35.5%; Score 621; DB 9; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-137;

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30462916.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: beget@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6066.X
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAF026ZE03_AF02422_1ac=6066.x

FEATURES

Location/Qualifiers
1..777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1041Y07"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 31.6%; Score 552.8; DB 4; Length 777;
Best Local Similarity 98.6%; Pred. No. 3.1e-121;
Matches 557; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

717 GATCTCCATCAGAGGCTCCATCTGTAGAGCAGCATGGAATGACTCCCTACTGAGGA 776
Db 4 GGGTCTCCTCATTAGGCTTCCATCTGTAGAGCAGCATGGAATGACTCCCTACTGAGGA 63
Qy 777 AGACTGCGCAGCAGCTGGGTTTCGAGAGTGTCTACCGGACACCGAGGTTGCCACAGG 836
Db 64 AGACTGCGCAGCAGCTGGGTTTCGAGAGTGTCTACCGGACACCGAGGTTGCCACAGG 123
Qy 837 ATTTTGGCAACAGCTTCTCAATCTTGAGATACAACTCCACATCCAGGAAAGCTCCACA 896
Db 124 ATTTTGGCAACAGCTTCTCAATCTTGAGATACAACTCCACATCCAGGAAAGCTCCACA 183
Qy 897 GATCGAATGCGCTTCCAGAGGATATCTCCCTCAGTGTCCACATGCGGAGCTGAAGG 956
Db 184 GATCGAATGCGCTTCCAGAGGATATCTCCCTCAGTGTCCACATGCGGAGCTGAAGG 243
Qy 957 CCATGACCGGCGGATCGTGGAGAGGGGCGTGCCTCGGATAGCAAGTGGCTTGGCAAG 1016
Db 244 CCATGACCGGCGGATCGTGGAGAGGGGCGTGCCTCGGATAGCAAGTGGCTTGGCAAG 303
Qy 1017 TGAATCTGCACTTGGGACCAACCATCTGTGAGAGGACGCTCATTTAGCGCCAGTGG 1076
Db 304 TGAATCTGCACTTGGGACCAACCATCTGTGAGAGGACGCTCATTTAGCGCCAGTGG 363
Qy 1077 TGCTGACCTGCGCCGACCTGCTTCTTGTGACCCCGGAGAAAGTCTTGAGAGGCTGGAAAG 1156
Db 364 TGCTGACCTGCGCCGACCTGCTTCTTGTGACCCCGGAGAAAGTCTTGAGAGGCTGGAAAG 423
Qy 1137 TGTAGCGCGGAGCAGCAGCACTGACAGAGTTGCTTGAGAGGAGCTTCATGCCAGATCA 1196
Db 424 TGTAGCGCGGAGCAGCAGCACTGACAGAGTTGCTTGAGAGGAGCTTCATGCCAGATCA 483
Qy 1197 TCATCAACAGCAATTACACCGATGAGAGAGAGTATGACATGCGCTCATGCGGCTGT 1256
Db 484 TCATCAACAGCAATTACACCGATGAGAGAGAGTATGACATGCGCTCATGCGGCTGT 543
Qy 1257 CCAAGCCCTGACCTGTCCGCTCA 1281
Db 544 CCAAGCCCTGACCTGTCCGCTCA 568

RESULT 13
DA830024
LOCUS
DEFINITION DA830024 PLAC1 Homo sapiens cDNA clone PLACE1004270 5', mRNA
sequence.
ACCESSION DA830024
VERSION DA830024.1 GI:82330405
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens (human)

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.

REFERENCE

1 (bases 1 to 852)
Kimura, K., Makamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Iwairant, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushiida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Matsushima, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560

CONTACT

Contact: Takao Isogai
Fuj Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: fuj-cdna@hri-fuj.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute (HRI); cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
HRI.

FEATURES

Location/Qualifiers
1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE104270"
/issue_type="Placenta"
/clone_lib="PLACE1"
/note="Vector: pME18FL3"

ORIGIN

Query Match 30.6%; Score 535.6; DB 9; Length 852;
Best Local Similarity 93.1%; Pred. No. 4.2e-117;
Matches 632; Conservative 0; Mismatches 22; Indels 25; Gaps 6;

1 CTCGAGACCTGAGAGGAGCAGCCACCGGAAATGCAATCTCCACAGAAACCTTTCAGC 60
Db 81 CTCGAGAGCACTGAGAGGAGCAGCCACCGGAAATGCAATCTCCACAGAAACCTTTCAGC 140
Qy 61 TGAAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 120
Db 141 TGAAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 200
Qy 121 CCAAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 180
Db 201 CCAAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 260
Qy 181 CCAAGCATCTCCAGCTGTACACCTCCAGGCGGGGCACTTCCAGCCG----- 228
Db 261 CCAAGCATCTCCAGCTGTACACCTCCAGGCGGGGCACTTCCAGCCGCGGATCTCCAGC 320
Qy 229 ---GGCATCTCCAGCCAGGATCTCCAGCCGGGGAATCTCCGGCTTGGCATCACTTC 285
Db 321 CCAAGCATCTCCAGCCAGGATCTCCAGCCGGGGAATCTCCGGCTTGGCATCACTTC 380

QY 286 CAGGTCCTCATCCGAGCAGTCATCCGACAGGTCAGCTCGGTGACAACTCCCAAC 345
 |||||
 Db 381 CAGGTCCTCATCCGAGCAGTCATCCGACAGGTCAGCTCGGTGACAACTCCCAAC 440
 |||||
 QY 346 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGC 405
 |||||
 Db 441 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGC 500
 |||||
 QY 406 CAGGTCAGACACAGCAACAGGGCCACAGGAGAGCCCAAGTACGAGCTGCCAAG 463
 |||||
 Db 501 CAGGTCAGACACAGCAACAGGGCCACAGGAGAGCCCAAGTACGAGCTGCCAAG 560
 |||||
 QY 464 TTCACTGCGGAGAGGAGCCAGAGCAGCTACCTCATCGGAGTGGCTCTCTCAT 522
 |||||
 Db 561 TTCACTGCGGAGAGGAGCCAGAGCAGCTACCTCATCGGAGTGGCTCTCTCAT 620
 |||||
 QY 523 TGCCCTGGTGTTCGCTCATCTCTTCAGTTCTGGAGGGCCACAGGAGATCAG 582
 |||||
 Db 621 TGCCCTGGTGTTCGCTCATCTCTTCAGTTCTGGAGGGCCACAGGAGATCAG 680
 |||||
 QY 583 GTACAGGAGCAGAGGAGAGCTGT---CCCAAGCAGCTGT---CGCTGACGGGG 635
 |||||
 Db 681 GTACAGGAGCAGAGGAGAGCTGTCCCAAGCAGCTGTTCGCTGATGAAAGGGGT 740
 |||||
 QY 636 TGGTGACTGCAAGCTGAA 654
 |||||
 Db 741 GGTGACTGCAAGCTGGA 759
 |||||

RESULT 14
 BX952562 640 bp mRNA linear EST 01-MAR-2004
 DEFINITION DKFZp781K14199.F1.781 (synonym: hlc64) Homo sapiens cDNA clone
 DKFZp781K14199.5, mRNA sequence.
 ACCESSION BX952562
 VERSION BX952562.1 GI:43429741
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 640)
 Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
 Robo,G., Han,M. and Wiemann,S.
 EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
 Unpublished (2003)
 COMMENT Contact: MIPS
 MIPS

REFERENCE Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZp781K14199) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcentzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 LOCATION/Qualifiers
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp781K14199"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_id="781 (synonym: hlc64)"
 /note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI;B;
 cDNA-collection"
 ORIGIN
 Query Match 30.5%; Score 532.8; DB 4; Length 640;

Best Local Similarity 97.0%; Pred. No. 1,86-116;
 Matches 559; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 CTCGAGACCATGAGAGGACACAGCCAGGAATGATCTCCACCAAGAACCTTCAGC 60
 |||||
 Db 65 CTCGAGACCATGAGAGGACACAGCCAGGAATGATCTCCACCAAGAACCTTCAGC 124
 |||||
 QY 61 TGAGCATCTCCAGCCAGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 120
 |||||
 Db 125 TGAGCATCTCCAGCCAGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 184
 |||||
 QY 121 CCAGGATCTCCAGCCAGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 180
 |||||
 Db 185 CCAGGATCTCCAGCCAGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 244
 |||||
 QY 181 CCAGGATCTCCAGCTGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 228
 |||||
 Db 245 CCAGGATCTCCAGCTGGATCTCCAGCTGGAGACCTCCAGCCGGGATCTCCAGC 304
 |||||
 QY 229 ---GGCATCTCCAGCCAGGATCTCCAGCCGGGATCTCCAGCTGGAGACCTTC 285
 |||||
 Db 305 CCAGGATCTCCAGCCAGGATCTCCAGCCGGGATCTCCAGCTGGAGACCTTC 364
 |||||
 QY 286 CAGGTCCTCATCCGAGCAGTCATCCGACAGGTCAGCTCGGTGACAACTCCCAAC 345
 |||||
 Db 365 CAGGTCCTCATCCGAGCAGTCATCCGACAGGTCAGCTCGGTGACAACTCCCAAC 424
 |||||
 QY 346 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGC 405
 |||||
 Db 425 CAGAGTGACTCTTTAGAGCAACACAGTGGGGCTGTACCATCCGATCATCTCTGC 484
 |||||
 QY 406 CAGGTCAGACACAGCAACAGGGCCACAGGAGAGCCAGGTCAGAGCTGCCAAGTT 465
 |||||
 Db 485 CAGGTCAGACACAGCAACAGGGCCACAGGAGAGCCAGGTCAGAGCTGCCAAGTT 544
 |||||
 QY 466 CACCTGGCGGAGGAGCCAGAGCAGTACCGCTCATCGGGTCGTCTCTCATATTGC 525
 |||||
 Db 545 CACCTGGCGGAGGAGCCAGAGCAGTACCGCTCATCGGGTCGTCTCTCATATTGC 604
 |||||
 QY 526 CTTGGTGTTCGCTCATCTCTTCAGTTCTG 561
 |||||
 Db 605 CTTGGTGTTCGCTCATCTCTTCAGTTCTG 640
 |||||

RESULT 15
 BP359643 579 bp mRNA linear EST 17-SEP-2004
 LOCUS BP359643 Sugano cDNA library, mammary gland T47D Homo sapiens cDNA
 clone TDR09260, mRNA sequence.
 DEFINITION BP359643
 ACCESSION BP359643
 VERSION BP359643.1 GI:52289656
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 579)
 Suzuki,Y., Yamashita,R., Shitota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 CONTACT: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Tel: 81-3-5449-5343
 Fax: 81-3-5449-5416
 Email: yusuzuki@ngc.jp.
 LOCATION/Qualifiers
 1..579

REFERENCE
 AUTHORS
 TITLE
 JOURNAL PUBMED
 COMMENT
 FEATURES
 source
 Query Match 30.5%; Score 532.8; DB 4; Length 640;

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TDR09260"
/tissue_type="mammary gland"
/cell_line="T47D"
/clone_lib="Sugano cDNA library, mammary gland T47D"
/note="mammary gland tumor"

```

/mol_type="mRNA"

```
/clone="TDR09260"
```

```
/cell_line="T47D"
```

```
/note="mammary gland tumor"
```

ity	28.24;	score 493.4;	DB 3;	length 3/3;
	99.84;	Pred. No. 5e-107;		

servative U; Mismatch

ORIGIN

Query Match	Score	DB	Length
28.2%	493.4	3	579

Best Local Similarity 99.8%; Pred. No. 5e-10/
Matches 494; Conservative 0; Mismatches 1; Indels 0
Gaps 0

1 CTCAGAGACATGGAGAGGACAGCCACGGGAATGATCTTCAGCAAGAACACCTTCAGC 60

Qy	1	CTCAGAGACCATGAGAGAGGAGCAGCCAGGAAATGATCTCCAGCAAGAAACACTTCAGC	60
Db	85	CTCAGAGACCATGAGAGAGGAGCAGCCAGGAAATGATCTCCAGCAAGAAACACTTCAGC	1444
Qy	61	TGAGCATCTCCAGGCCAGGCATCTTCAGCTGGGACACTTCGAGGCCGGGCATCTCCAGC	120
Db	145	TGAGGATCTCCAGGCCAGGCATCTTCAGCTGGGACACTTCGAGGCCGGGCATCTCCAGC	204
Qy	121	CCAGGCATCTCCAGGCCAGGCATCTTCAGCTGGGACACTTCGAGGCCGGGCATCTCCAGC	180
Db	205	CCAGGCATCTCCAGGCCAGGCATCTTCAGCTGGGACACTTCGAGGCCGGGCATCTCCAGC	264
Qy	181	CCAGGCATCTCCAGCTGGTACACTTCGAGGCCGGGCATCTTCGAGGCCGGGCATCTCCAGC	240
Db	265	CCAGGCATCTCCAGCTGGTACACTTCGAGGCCGGGCATCTTCGAGGCCGGGCATCTCCAGC	324
Qy	241	CCAGGCATCTCCAGGCCGGGCCATCTTCGAGCTTCGAGCATCACTTTCGAGGCTTCATCCGG	300
Db	325	CCAGGCATCTCCAGGCCGGGCCATCTTCGAGCTTCGAGCATCACTTTCGAGGCTTCATCCGG	384
Qy	301	CAGGTCAATCATCCGCGAGGTCAAGCTCCGGTGAACAACCTCCCAACCAAGATGACTTGT	360
Db	385	CAGGTCAATCATCCGCGAGGTCAAGCTCCGGTGAACAACCTCCCAACCAAGATGACTTGT	444
Qy	361	TAAAGCAACACAGATGGGGGGCTGTACCAATCCGATCAATCTCTCGCAGGTCAAGACAGC	420
Db	445	TAAAGCAACACAGATGGGGGGCTGTACCAATCCGATCAATCTCTCGCAGGTCAAGACAGC	504
Qy	421	AACAGGGCCACCAAGGAGAGGCCCAAGGTACGAGCTTCGCAAGTTCACTTGGCGGAGGG	480
Db	505	AACAGGGCCACCAAGGAGAGGCCCAAGGTACGAGCTTCGCAAGTTCACTTGGCGGAGGG	564
Qy	481	CCGAAGCAGACTAC 495	
Db	565	CCGAAGCAGACTAC 579	

Search completed: September 18, 2006, 10:30:20
Job time : 8725 secs

Search completed: September 18, 2006, 10:30:20
Job time : 8725 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 07:36:49 / Search time 1148 Seconds
(without alignments)
10616.281 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 1748
Sequence: 1 ctcagagacatcgagagag99.....ggtctgtgactcagagaa 1748

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_8:*

1: geneseqn1980s:*

2: geneseqn1980s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

14: geneseqn2005s:*

15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1748	100.0	1748	6	AAD27734 Human tra
2	1689	96.6	1707	12	ADH17439 Human NOV
3	1689	96.6	1708	12	ADH17457 Human NOV
4	1687.4	96.5	1689	4	AAD05796 Human tra
5	1687.4	96.5	1689	10	ADH10392 Human cel
6	1687.4	96.5	1689	12	ADH146916 Human tra
7	1687.4	96.5	1689	14	ADH150145 Human end
8	1670.4	95.6	2380	14	ADH275551 Human mos
9	1670.4	95.6	2393	10	ADH31742 Human 291
10	1663.4	95.2	2067	4	AAD05797 Human tra
11	1663.4	95.2	2067	10	ADH10394 Human cel
12	1663.4	95.2	2067	12	ADH146918 Human tra
13	1663.4	95.2	2067	14	ADH150053 Human end
14	1645.4	94.1	2432	12	ADH17413 Human NOV
15	1643.8	94.0	2432	12	ADH17453 Human NOV
16	1643.8	94.0	2432	12	ADH17451 Human NOV
17	1643.8	94.0	2432	12	ADH17449 Human NOV
18	1640	93.8	1771	12	ADH17433 Human NOV

19	1640	93.8	1779	12	ADH17429	Adh17429 Human NOV
20	1479	84.6	2310	12	ADH17437	Adh17437 Human NOV
21	1429	81.8	1614	6	ABK31788	Abk31788 DNA encod
22	1402	80.2	1671	14	AEA19673	Aea19673 Novel hum
23	1369	78.3	1626	12	ADH17431	Adh17431 Human NOV
24	1302.8	74.5	1314	4	AAf83971	Aaf83971 Nucleotid
25	1302.8	74.5	1314	4	AAf83971	Aaf83971 Human SER
26	1302.8	74.5	1314	12	ADH17441	Adh17441 Human NOV
27	1299.6	74.3	1314	10	ADH29369	Adh29369 Human ser
28	1278.2	73.1	1314	10	ADH29366	Adh29366 Human CDN
29	1277.8	73.1	1574	6	AAD0575	Aad0575 Human pro
30	1258.6	72.0	1434	8	ABZ22864	Abz22864 Human den
31	1252.8	71.7	2192	4	AAH14850	Aah14850 Human CDN
32	1155.2	66.1	1409	13	ACN42023	Acn42023 Human pro
33	1147	65.6	3103	13	ACN42023	Acn42023 Human dia
34	1143.4	65.4	1146	12	ADH17443	Adh17443 Human NOV
35	1131.8	64.7	1407	10	ADH19039	Adh19039 Human pro
36	1120.4	64.1	1218	12	ADH17435	Adh17435 Human NOV
37	1120.4	64.1	1218	12	ADH17447	Adh17447 Human NOV
38	1120	64.1	1448	10	ADH79040	Adh79040 Human pro
39	1105	63.2	1203	12	ADH17421	Adh17421 Human NOV
40	1103.4	63.1	1203	12	ADH17427	Adh17427 Human NOV
41	1101.8	63.0	1203	12	ADH17425	Adh17425 Human NOV
42	1068.4	61.1	1078	4	AAf83972	Aaf83972 Nucleotid
43	1062.6	60.8	1341	8	ABZ22865	Abz22865 Human den
44	1053	60.2	1230	6	AAD27743	Aad27743 Human tra
45	1050	60.1	1077	10	ADH29368	Adh29368 Human ser

ALIGNMENTS

RESULT 1	
AAD27734	standard; DNA; 1748 BP.
ID	AAD27734
XX	
AC	AAD27734;
XX	
DT	18-APR-2002 (first entry)
XX	
DE	Human transmembrane serine protease DNA.
XX	
KW	Human; transmembrane serine protease; gene therapy; metastasis; tumour;
KW	chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;
KW	atherosclerosis; neurodegenerative disease; neuroprotective; cytoskeletal;
KW	pathogenic infection; antiinflammatory; antitreticlerotic;
XX	antibacterial; ds.
XX	
OS	Homo sapiens.
XX	
FT	
FT	Key
FT	11.1699
FT	/tag= a
FT	/product= "Human transmembrane serine protease"
XX	
PN	WO200196538-A2.
XX	
PD	20-DEC-2001.
XX	
PF	12-JUN-2001; 2001WO-EP06618.
XX	
PR	13-JUN-2000; 2000US-0211234P.
PR	13-APR-2001; 2001US-0283533P.
PR	16-APR-2001; 2001US-0283648P.
XX	
PA	(FARB) BAYER AG.
XX	
PI	Xiao Y, Gedlich R;
XX	WPI, 2002-098065/13.
DR	P-PSDS; AAE17238.
XX	
PT	Novel isolated polynucleotide encoding transmembrane serine protease

PT polypeptide, for treating chronic obstructive pulmonary disease, tumor
PT angiogenesis, inflammation, atherosclerosis and neurodegenerative
PT disease.

PS Claim 1; Page 111-112; 120pp; English.

CC The present invention relates to an isolated polynucleotide encoding a
CC transmembrane serine protease polypeptide. Transmembrane serine protease
CC gene is useful in gene therapy. The invention also relates to a
CC pharmaceutical composition which is useful for modulating the activity of
CC transmembrane serine protease in a disease, such as chronic obstructive
CC pulmonary disease (COPD), metastasis of malignant cells, tumor
CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease
CC or pathogenic infection. Transmembrane serine protease is useful as a
CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is
CC useful for generating antibodies against it and in various assay systems.
CC The present sequence is a human transmembrane serine protease DNA

XX Sequence 1748 BP; 379 A; 550 C; 487 G; 332 T; 0 U; 0 Other;

Query Match 100.0%; Score 1748; DB 6; Length 1748;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTCAGAGACATGAGAGAGACACCCAGGAAATGATCTTCACAGCAAGAACCTTCAGC	60
DB	1	CTCAGAGACATGAGAGAGAGACACCCAGGAAATGATCTTCACAGCAAGAACCTTCAGC	60
QY	61	TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTCCAGC	120
DB	61	TGGAGCATCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTCCAGC	120
QY	121	CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTCCAGC	180
DB	121	CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTCCAGC	180
QY	181	CCAGGCACTCTCCAGCTGGATCACTTCACAGCCGGGATCTTCAGAGCCGGGATCTTCAGC	240
DB	181	CCAGGCACTCTCCAGCTGGATCACTTCACAGCCGGGATCTTCAGAGCCGGGATCTTCAGC	240
QY	241	CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCAGC	300
DB	241	CCAGGCACTCTCCAGCCCGAGGATCTCCAGCTGGAGACACTTCAGAGCCGGGATCTTCAGC	300
QY	301	CAGGTCATCATCCGCCAGGTGAGCTCGGTGACAACTCCCAACAGAGGTGACCTTGT	360
DB	301	CAGGTCATCATCCGCCAGGTGAGCTCGGTGACAACTCCCAACAGAGGTGACCTTGT	360
QY	361	TAGAGCAACACAGTGGGGGCTGTACCATCGATCATCTCTCGCAGAGTACAGACAGC	420
DB	361	TAGAGCAACACAGTGGGGGCTGTACCATCGATCATCTCTCGCAGAGTACAGACAGC	420
QY	421	AACCAAGGCGCCAGGAGAGACCCAGGTGACAGAGCTGCCAAAGTTTCACTGGCCGAGG	480
DB	421	AACCAAGGCGCCAGGAGAGACCCAGGTGACAGAGCTGCCAAAGTTTCACTGGCCGAGG	480
QY	481	CCAGAAACAGGTACCGCTCATCGGAGTGGTCTCTCTCATTTGCCCTGGTGTTCCT	540
DB	481	CCAGAAACAGGTACCGCTCATCGGAGTGGTCTCTCTCATTTGCCCTGGTGTTCCT	540
QY	541	CATCATCTCTTTCAGATTGAGAGGCGACACAGGATCAAGGTACAAAGAGAGAGGA	600
DB	541	CATCATCTCTTTCAGATTGAGAGGCGACACAGGATCAAGGTACAAAGAGAGAGGA	600
QY	601	GAGCTGTCCCAAGACGCTGTTCCTGTGACGGGGTGTGACGTGCAAGCTGAAAGAGGA	660
DB	601	GAGCTGTCCCAAGACGCTGTTCCTGTGACGGGGTGTGACGTGCAAGCTGAAAGAGGA	660
QY	661	CGAGCTGGGCTGCGTGAAGTTTGAATGAGCAAGTCTCTGTTAAATCTACTCTGGGTC	720
DB	661	CGAGCTGGGCTGCGTGAAGTTTGAATGAGCAAGTCTCTGTTAAATCTACTCTGGGTC	720
QY	721	CTCCATCATGAGCTTCCCATCTGTAGACAGCACTGAATGACTCTTACTGAGAAAGAC	780

DB	721	CTCCATCATGAGCTTCCCATCTGTAGACAGCACTGAATGACTCTTACTGAGAAAGAC	780
QY	781	CTGCCAGACGTGGGCTTTCAGAGGTGCTCAACCGACAAACGAGGTTCCCAAGGATTT	840
DB	781	CTGCCAGACGTGGGCTTTCAGAGGTGCTCAACCGACAAACGAGGTTCCCAAGGATTT	840
QY	841	TGCCAAGAGCTTCTCAATCTTGAATCAACTCCACCATCCAGAGAAAGCTTCACAGGTC	900
DB	841	TGCCAAGAGCTTCTCAATCTTGAATCAACTCCACCATCCAGAGAAAGCTTCACAGGTC	900
QY	901	TGAATGCCCTTCCAGCGGTATATCTCCCTCCAGTGTTCCTCCAGCTGGAGTGGCCAT	960
DB	901	TGAATGCCCTTCCAGCGGTATATCTCCCTCCAGTGTTCCTCCAGCTGGAGTGGCCAT	960
QY	961	GACCGGCGGATCGTGGAGAGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG	1020
DB	961	GACCGGCGGATCGTGGAGAGGGGCGTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG	1020
QY	1021	TCTGCACTTGGGCAACCAACCATCTGTGAGAGGACGCTCATTTGACCCAGTGGTGTCT	1080
DB	1021	TCTGCACTTGGGCAACCAACCATCTGTGAGAGGACGCTCATTTGACCCAGTGGTGTCT	1080
QY	1081	CATGCGCGCCACTGCTTCTTCGTGACCCGGAGAGAGTCTGTGAGGGCTTGAAGGTGTA	1140
DB	1081	CATGCGCGCCACTGCTTCTTCGTGACCCGGAGAGAGTCTGTGAGGGCTTGAAGGTGTA	1140
QY	1141	CGGGGGACAGCAACCTGACACCGATGTCCTGAGAGGAGCCCTCCATTGCCAGATCATCAT	1200
DB	1141	CGGGGGACAGCAACCTGACACCGATGTCCTGAGAGGAGCCCTCCATTGCCAGATCATCAT	1200
QY	1201	CAACAGCAATTACACCGATGAGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCCAA	1260
DB	1201	CAACAGCAATTACACCGATGAGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCCAA	1260
QY	1261	GCCCTGACCTTGTGCTGCTCATTCACCTGCTGCTTCCCTCCATGCAATGAGACAGCTT	1320
DB	1261	GCCCTGACCTTGTGCTGCTCATTCACCTGCTGCTTCCCTCCATGCAATGAGACAGCTT	1320
QY	1321	TAGCTCAATGAGACCTGCTGATGACAGGCTTTGGCAAGCCAGGGAGACAGATGACAA	1380
DB	1321	TAGCTCAATGAGACCTGCTGATGACAGGCTTTGGCAAGCCAGGGAGACAGATGACAA	1380
QY	1381	GACATCCCTCTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA	1440
DB	1381	GACATCCCTCTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA	1440
QY	1441	CTACTTGGCTATGACAGTTAAGTTTACCTTACCCCAAGATGATGTGCTGGGACCTTGTGG	1500
DB	1441	CTACTTGGCTATGACAGTTAAGTTTACCTTACCCCAAGATGATGTGCTGGGACCTTGTGG	1500
QY	1501	GGGAGAGACCTCCGACAGGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACCG	1560
DB	1501	GGGAGAGACCTCCGACAGGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACCG	1560
QY	1561	CTGGTACCTGAGAGGTGACACAGCTGGGACAGGCTGTGGCCAGAGAAACAACTGG	1620
DB	1561	CTGGTACCTGAGAGGTGACACAGCTGGGACAGGCTGTGGCCAGAGAAACAACTGG	1620
QY	1621	TGTGTACACCAAGTGAAGAGTTTCTTCCCTGGAATTTACAGCAAGATGAGAGAGAGGT	1680
DB	1621	TGTGTACACCAAGTGAAGAGTTTCTTCCCTGGAATTTACAGCAAGATGAGAGAGAGGT	1680
QY	1681	GCGATTGAGAAATCTTAACAGCTGGCTGTGCTGTGACAGACACCGGCTGTGTGAC	1740
DB	1681	GCGATTGAGAAATCTTAACAGCTGGCTGTGCTGTGACAGACACCGGCTGTGTGAC	1740
QY	1741	TCGAGAAA 1748	
DB	1741	TCGAGAAA 1748	

RESULT 2

Db 670 GGCTGCGTAGGTTGACTGGGACAAGCTCTGCTTAAATCTACTGCGGCTCTCCCAT 729
Qy 728 CAGTGGTCTCCATCTCTAGACGAACCTGGAATGATCTCTAATGAGAAAGCTTGCAG 787
Db 730 CAGTGGTCTCCATCTCTAGACGAACCTGGAATGATCTCTAATGAGAAAGCTTGCAG 789
Qy 788 CAGTGGGTTTCGAGATGCTCACCGGACAACGAGGTTGGCCACAGGAAATTTTGGCAAC 847
Db 790 CAGTGGGTTTCGAGATGCTCACCGGACAACGAGGTTGGCCACAGGAAATTTTGGCAAC 849
Qy 848 AGCTTCTCAATCTTGAATACAACTCCACATCCAGAAAGCTTCCACAGTCTGAATGC 907
Db 850 AGCTTCTCAATCTTGAATACAACTCCACATCCAGAAAGCTTCCACAGTCTGAATGC 909
Qy 908 CCTTCCAGAGGGATATCTCTCCAGGTTCCACATGCGGACGAGGAGGAGGAGGAGGAGG 967
Db 910 CCTTCCAGAGGGATATCTCTCCAGGTTCCACATGCGGACGAGGAGGAGGAGGAGGAGG 969
Qy 968 CGGATCGTGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCAC 1027
Db 970 CGGATCGTGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCAC 1029
Qy 1028 TTGCGGACACCCACATCTGTGAGGACGCTCATTTGACGCCCAAGTGGTCTCACTGCC 1087
Db 1030 TTGCGGACACCCACATCTGTGAGGACGCTCATTTGACGCCCAAGTGGTCTCACTGCC 1089
Qy 1088 GCCCACTGCTTCTCGTGAACCGGGAGAAAGTCTGGAGGGCTGGAAGGTATAGCCGGAGC 1147
Db 1090 GCCCACTGCTTCTCGTGAACCGGGAGAAAGTCTGGAGGGCTGGAAGGTATAGCCGGAGC 1149
Qy 1148 ACCAGCAACCTGACCAAGTTGCTGAGGACAGCTCATTTGCGGATCATCATCAAGAC 1207
Db 1150 ACCAGCAACCTGACCAAGTTGCTGAGGACAGCTCATTTGCGGATCATCATCAAGAC 1209
Qy 1208 AATTACACCGATGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
Db 1210 AATTACACCGATGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
Qy 1268 ACCGTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
Db 1270 ACCGTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
Qy 1328 AATGAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
Db 1330 AATGAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
Qy 1388 CCTTCTCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1447
Db 1390 CCTTCTCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
Qy 1448 GTCTATGACGTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1507
Db 1450 GTCTATGACGTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1509
Qy 1508 GACTCTGCGGAGGAGACAGCGGGGGGCTCTTCTGTGTGAGCAAGAACCGCTGTGTAC 1567
Db 1510 GACTCTGCGGAGGAGACAGCGGGGGGCTCTTCTGTGTGAGCAAGAACCGCTGTGTAC 1569
Qy 1568 CTGGCAGGTGTACCAAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACTGTGTGTAC 1627
Db 1570 CTGGCAGGTGTACCAAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACTGTGTGTAC 1629
Qy 1628 ACCAAAGTACAGAAAGTTCTTCCCGATTATACAGCAAGATGAGAGGAGGAGTGCATTC 1687
Db 1630 ACCAAAGTACAGAAAGTTCTTCCCGATTATACAGCAAGATGAGAGGAGGAGTGCATTC 1689
Qy 1688 AGAAATCC 1696
Db 1690 AGAAATCC 1698

RESULT 3

ADH17457/C
ID ADH17457 standard; cDNA; 1708 BP.
XX
AC ADH17457;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human NOV13b cDNA - SEQ ID 147.
XX
KW NOX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW protozoal; neurotropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW dermatological; antiepileptic; antiparkinsonian; metabolic; diabetes;
KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
KW epilepsy; immune; osteoarthritis; haemopoietic;
KW inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
KW cell differentiation; proliferation; haemopoiesis; wound healing;
KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
KW pharmacogenomic; human; ss; gene.
OS Homo sapiens.
XX
PN WO2003093432-A2.
XX
PD 13-NOV-2003.
XX
PE 02-MAY-2003; 2003MO-US013690.
XX
PR 02-MAY-2002; 2002US-0377321P.
PR 08-MAY-2002; 2002US-0378730P.
PR 24-MAY-2002; 2002US-0383075P.
PR 29-MAY-2002; 2002US-0384044P.
PR 30-MAY-2002; 2002US-0384215P.
PR 30-MAY-2002; 2002US-0384226P.
PR 30-MAY-2002; 2002US-0384297P.
PR 30-MAY-2002; 2002US-0384337P.
PR 30-MAY-2002; 2002US-0384352P.
PR 31-MAY-2002; 2002US-0385211P.
PR 02-JUL-2002; 2002US-0393213P.
PR 09-AUG-2002; 2002US-0402154P.
PR 09-AUG-2002; 2002US-0402171P.
PR 09-AUG-2002; 2002US-0402204P.
PR 09-AUG-2002; 2002US-0402205P.
PR 22-AUG-2002; 2002US-0405175P.
PR 27-AUG-2002; 2002US-0406129P.
PR 23-SEP-2002; 2002US-0412954P.
PR 30-SEP-2002; 2002US-0414975P.
PR 07-OCT-2002; 2002US-0416611P.
PR 24-OCT-2002; 2002US-0420851P.
PR 31-OCT-2002; 2002US-0422547P.
PR 01-MAY-2003; 2003US-00428275.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alvarez E, Anderson DW, Boldog FL, Catterton E, Edinger SR;
PI Fernandes ER, Gerlach VU, Gorman L, Grose WM, Guo X, Ji W;
PI Kehuda R, Li L, Macdougall JR, Padigaru M, Paturajan M;
PI Verneton JD, Rastelli L, Shinkens RA, Sytek KA, Stone DJ;
PI Vernet CAM, Voss EZ, Zhong W;
XX
DR MPI; 2004-053040/05.
DR P-PSDB; ADH17458.
XX
PT New isolated NOX polypeptide, useful for preventing, diagnosing or
PT treating NOX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
PS Claim 20; SEQ ID NO 147; 478pp; English.
XX
CC The invention relates to a novel isolated NOX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,

CC cardiant, hypotensive, antiarteriosclerotic, anorectic, virucide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
CC antiinflammatory, dermatological, antisthmatic and antidiabetic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX cDNA of the invention.

XX Sequence 1708 BP; 323 A; 477 C; 539 G; 369 T; 0 U; 0 Other;

Query Match 96.6%; Score 1689; DB 12; Length 1708;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACCATGAGAGGAGCAGACCCAGGAGATGCACTTCACAGAACACCTTCAGCTGAGCA 67
DB 1698 ACCATGAGAGGAGCAGACCCAGGAGATGCACTTCACAGAACACCTTCAGCTGAGCA 1639
QY 68 TCTCCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGGCGGAGATCTCCAGCCAGCA 127
DB 1638 TCTCCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGGCGGAGATCTCCAGCCAGCA 1579
QY 128 TCTCCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGGCGGAGATCTCCAGCCAGCA 187
DB 1578 TCTCCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGGCGGAGATCTCCAGCCAGCA 1519
QY 188 TCTCCAGCTGGATACCTCCAGGCGGAGATCTCCAGGCGGAGATCTCCAGCCAGCA 247
DB 1518 TCTCCAGCTGGATACCTCCAGGCGGAGATCTCCAGGCGGAGATCTCCAGCCAGCA 1459
QY 248 TCTCCAGCCCGGAGATCTCCGAGCTCTGAGCACTTCAGAGTCTTCATCCGAGAGTCA 307
DB 1458 TCTCCAGCCCGGAGATCTCCGAGCTCTGAGCACTTCAGAGTCTTCATCCGAGAGTCA 1399
QY 308 TCAATCCGAGGATCAAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTAAGCA 367
DB 1398 TCAATCCGAGGATCAAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTAAGCA 1339
QY 368 ACACCAATGGGGGCTGATACCATCGATCTCTGCAAGTCAAGCAGCAACAGG 427
DB 1338 ACACCAATGGGGGCTGATACCATCGATCTCTGCAAGTCAAGCAGCAACAGG 1279
QY 428 GCCACCAAGGAGAGCCAGGATGAGAGCTTCGCAAGTTCACTGAGGAGGAGCAAG 487
DB 1278 GCCACCAAGGAGAGCCAGGATGAGAGCTTCGCAAGTTCACTGAGGAGGAGCAAG 1219
QY 488 CAGGTACCGCTCATCGGGTGGTCTCTCTCATTTGCCCTGGTGGTTTGGCTCATATC 547
DB 1218 CAGGTACCGCTCATCGGGTGGTCTCTCTCATTTGCCCTGGTGGTTTGGCTCATATC 1159
QY 548 CTCTTCCAGTTCTGGAGGAGCAGAGGATCAAGTCAAGAGAGAGGAGAGCTGT 607
DB 1158 CTCTTCCAGTTCTGGAGGAGCAGAGGATCAAGTCAAGAGAGAGGAGAGCTGT 1039
QY 608 CCCAAGCAGCTGTTCTGCTGAGCGGGTGGTGACTGCAAGTGAAGAGTGAAGAGCTG 667
DB 1098 CCCAAGCAGCTGTTCTGCTGAGCGGGTGGTGACTGCAAGTGAAGAGTGAAGAGCTG 1039
QY 668 GGCTGCGTGAAGTTGA CTGGAGACAAGTCTGTGTTAAATCTACTGAGGCTCTCCAT 727

DB 1038 GGCTGCGTGAAGTTGACTGGAGCAAGTCTGTGTTAAATCTACTGAGGCTCTCCAT 979
QY 728 CAGTGGCTTCCCATCTGTACAGCACTGGAATGACTCTTACTCAGAGAGACTGGCAG 787
DB 978 CAGTGGCTTCCCATCTGTACAGCACTGGAATGACTCTTACTCAGAGAGACTGGCAG 919
QY 788 CAGTGGGTTTCCAGAGTGTCAACCGGACAAACGAGGTTGCCACAGGATTTTGGCAAC 847
DB 918 CAGTGGGTTTCCAGAGTGTCTACCGGACAAACGAGGTTGCCACAGGATTTTGGCAAC 859
QY 848 AGCTTCTCAATCTTGAATCAACTCCACCATCCAGAGAAAGCTTCCACAGTCTGAATGC 907
DB 858 AGCTTCTCAATCTTGAATCAACTCCACCATCCAGAGAAAGCTTCCACAGTCTGAATGC 799
QY 908 CTTTCCAGGAGTATATCTCTCTCCAGTGTTCCTCACTGCGGACTGAGGAGCCTACCGG 967
DB 798 CTTTCCAGGAGTATATCTCTCTCCAGTGTTCCTCACTGCGGACTGAGGAGCCTACCGG 739
QY 968 CGGATCGTGGAGGAGGAGGCTGGCTGGATGAGCAAGTGGCTTGGCAAGTGAATCTGCAC 1027
DB 738 CGGATCGTGGAGGAGGAGGCTGGCTGGATGAGCAAGTGGCTTGGCAAGTGAATCTGCAC 679
QY 1028 TTGGGACACCAACCAATCTGTGAGAGGACGCTCATTTGACGCCAGTGGTGTCTCACTGCC 1087
DB 678 TTGGGACACCAACCAATCTGTGAGAGGACGCTCATTTGACGCCAGTGGTGTCTCACTGCC 619
QY 1088 GCCCATCTGCTTCTTCTGAGACCCGGAGAGAGTCTTGGAGGCTGGAAAGTGTACCGGGC 1147
DB 618 GCCCATCTGCTTCTTCTGAGACCCGGAGAGAGTCTTGGAGGCTGGAAAGTGTACCGGGC 559
QY 1148 ACCAGCAACCTGACCAAGTGGCTGTGAGGAGGCTTCCATTCGAGATCATCATCAAGAGC 1207
DB 558 ACCAGCAACCTGACCAAGTGGCTGTGAGGAGGCTTCCATTCGAGATCATCATCAAGAGC 499
QY 1208 AATTACACCGATGAGAGAGAGCACTATGATGAGCTGAGCTTACGAGCTGTCCAAAGCCTG 1267
DB 498 AATTACACCGATGAGAGAGAGCACTATGATGAGCTGAGCTTACGAGCTGTCCAAAGCCTG 439
QY 1268 ACCCTGTGCTGCATCATCAACCTGTCTGCTCCCATGATGAGACAGACCTTTAGCCTC 1327
DB 438 ACCCTGTGCTGCATCATCAACCTGTCTGCTCCCATGATGAGACAGACCTTTAGCCTC 379
QY 1328 AATGAGACTGTGTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAAGATCC 1387
DB 378 AATGAGACTGTGTGATCAAGGCTTTGGCAAGACAGGAGACAGATGACAAAGATCC 319
QY 1388 CCGTCTCCGAGAGTGTGAGGATCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG 1447
DB 318 CCGTCTCCGAGAGTGTGAGGATCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG 259
QY 1448 GTCTATGACAGTTACCTTACCCCAAGGATGATGTGTCTGGGAGCTTCTGTGGGAGGAGA 1507
DB 258 GTCTATGACAGTTACCTTACCCCAAGGATGATGTGTCTGGGAGCTTCTGTGGGAGGAGA 199
QY 1508 GACTCTCTCCAGGAGAGCACCGGGGGCTCTTGTCTGTGACGAGAACACCGCTGTATC 1567
DB 198 GACTCTCTCCAGGAGAGCACCGGGGGCTCTTGTCTGTGACGAGAACACCGCTGTATC 139
QY 1568 CTGGCAGGTGTCAACAGCTGGGAGCAGGCTGTGGCAGAGAGAAACAACTGTGTGTAC 1627
DB 138 CTGGCAGGTGTCAACAGCTGGGAGCAGGCTGTGGCAGAGAGAGAAACAACTGTGTGTAC 79
QY 1628 ACCAAAGTGAACAGAGTTCTTCCCTGATTTTACGAAATGAGAGACGAGTGCATTC 1687
DB 78 ACCAAAGTGAACAGAGTTCTTCCCTGATTTTACGAAATGAGAGACGAGTGCATTC 19
QY 1688 AGAAAAATCC 1696
DB 18 AGAAAAATCC 10

RESULT 4

AAD05796
 ID AAD05796 standard; DNA; 1689 BP.
 XX
 AC AAD05796;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transmembrane serine protease (Endothelinase 2-S) DNA.
 XX
 KW Human: endothelinase 2-S; protease domain; cytostatic; vulnery; wound;
 KW angiogenesis; cardiovascular disease; psoriasis; neovascular disease;
 KW chronic inflammatory disease; ocular disorder; circulatory disorder;
 KW crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KW liver cirrhosis; osteoarthritis; systemic sclerosis; oesophageal;
 KW inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
 KW systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
 KW transmembrane serine protease; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 1..1689
 FT /tag= a
 FT /product= "Human endothelinase 2-S protein"
 XX
 PN M0200136604-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000MO-US031803.
 XX
 PR 18-NOV-1999; 99US-016391P.
 PR 22-SEP-2000; 2000US-0234840P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Madison BL, Ong EO;
 XX
 DR WPI: 2001-336001/35.
 DR P-PSDB: AAE01943.
 XX
 PT New nucleic acid encoding a protein comprising endothelinase activly
 PT useful in the prevention and treatment of e.g. vascular malformations,
 PT cardiovascular disorders, and chronic inflammatory disease.
 XX
 PS Example 2; Page 135-137; 152pp; English.
 CC The present DNA sequence encodes human short form transmembrane serine
 CC protease (Endothelinase 2-S) protein. The invention relates to an
 CC endothelinase protein, endothelinase protease domain and their
 CC corresponding nucleic acid molecules. An endothelinase protein or protease
 CC domain of it is useful for the treatment and diagnosis of disorders
 CC associated with aberrant angiogenesis or undesired neovascularisation.
 CC The undesired angiogenesis is associated with disorders selected from
 CC solid neoplasm, vascular malformations and cardiovascular disorders such
 CC as angiodioma, angiolipoma, atherosclerosis, restenosis/reperfusion
 CC injury, arteriovenous malformations, haemangiomas and vascular
 CC adhesions, dyschondroplasia with vascular hamartomas (Pafucel's
 CC syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
 CC syndrome) and Von Hippel Lindau syndrome, chronic inflammatory bowel
 CC disease, nonhealing fractures, haemophilic joints, psoriasis, rheumatoid
 CC arthritis, venous stasis ulcers, granuloma-burns, hypertrophic scars,
 CC liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
 CC granuloma and systemic sclerosis and aberrant wound repair, circulatory
 CC disorders Raynaud's phenomenon, crest syndromes such as calcinosis,
 CC oesophageal, dyomeolity, sclerodactyly and teangiectasis, dermatological
 CC disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
 CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
 CC stain, blue rubber blab nevus syndrome, Klippel-Trenaunay-Weber syndrome
 CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
 CC caused by ocular neovascular disease, corneal graft neovascularisation,

CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia
 CC and corneal neovascularisation. The nucleic acids of the invention are
 CC also used in gene therapy. The invention also provides method for
 CC screening compounds that modulate angiogenesis
 XX
 SO Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;
 Query Match 96.5%; Score 1687.4; DB 4; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 11 ATGAGAGGACACGACCGGGAATGATCTCCAGACAGAAACCTTACCTGAGCATCT 70
 1 ATGAGAGGACACGACCGGGAATGATCTCCAGACAGAAACCTTACCTGAGCATCT 60
 71 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCT 130
 61 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCT 120
 131 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCT 190
 121 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCT 180
 191 CCAAGCTGTACACTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGCCAGGACATCT 250
 181 CCAAGCTGTACACTCCAGGCGGGCATCTCCAGGCGGGCATCTCCAGCCAGGACATCT 240
 251 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCTCA 310
 241 CCAAGCCAGGACATCTCCAGCTGGGACACCTCCAGGCGGGCATCTCCAGCCAGGACATCTCA 300
 311 TCCGACAGGTCAAGCTCGGTGACCAACCTCCCAACAGAGTGTACCTTTAGAACAA 370
 301 TCCGACAGGTCAAGCTCGGTGACCAACCTCCCAACAGAGTGTACCTTTAGAACAA 360
 371 CCAAGTGGGGCTGTACCATTCCTATCTCTGACAGGTTCAGACCAAGAGGCC 430
 361 CCAAGTGGGGCTGTACCATTCCTATCTCTGACAGGTTCAGACCAAGAGGCC 420
 431 ACCAGGAGAGCCAGGATACAGGCTGCCAAGTTCACTGACGGGAGGACAGAAACAG 490
 421 ACCAGGAGAGCCAGGATACAGGCTGCCAAGTTCACTGACGGGAGGACAGAAACAG 480
 491 CTACCGCTCATCGGGTGCCTGCTCTCTCTCATATGCTGATCATCTCTC 550
 481 CTACCGCTCATCGGGTGCCTGCTCTCTCTCATATGCTGATCATCTCTC 540
 551 TTCCAGTTCTGACAGGACCAACAGGATACAGTGTACAGAGAGAGAGAGAGTGTCCC 610
 541 TTCCAGTTCTGACAGGACCAACAGGATACAGTGTACAGAGAGAGAGAGTGTCCC 600
 611 AAGCAGCCTGTTGCTGTGACAGGAGGTGAGTCAAGCTGAAGAGTGAAGAGCTGGC 670
 601 AAGCAGCCTGTTGCTGTGACAGGAGGTGAGTCAAGCTGAAGAGTGAAGAGCTGGC 660
 671 TGCCTGAGGTTTGAAGTGGGACAAAGTCTCTGCTTAAATTAATCTGCTGCTCCATCAG 730
 661 TGCCTGAGGTTTGAAGTGGGACAAAGTCTCTGCTTAAATTAATCTGCTGCTCCATCAG 720
 731 TGGCTTCCCATCTGTAGACCAACTGGAATGATCTCTCACTCAAGAGAGACTCTGCACAG 790
 721 TGGCTTCCCATCTGTAGACCAACTGGAATGATCTCTCACTCAAGAGAGACTCTGCACAG 780
 791 CTGGGTTTCAGAGATGCTCACCGGACCAAGAGGTGCCACAGGAGATTTTGCACAGC 850
 781 CTGGGTTTCAGAGATGCTCACCGGACCAAGAGGTGCCACAGGAGATTTTGCACAGC 840
 851 TTCTCAATCTTGAGATCAACTCCACCATTCAGAGAAAGCTTCCACAGTGTGAATGCTCT 910
 841 TTCTCAATCTTGAGATCAACTCCACCATTCAGAGAAAGCTTCCACAGTGTGAATGCTCT 900
 911 TCCAGGAGATATCTCCCTCCAGTGTTCCTCCACTGCGGAGCTGAGAGGACATGACCGGGCGG 970

Db 901 TCCACGGGTATATCTCTCCAGTGTTCACATGCGGAACTGAGGCCATGACCGGGCGG 960
 Qy 971 ATGTGGAGAGGGGGCTGGCTCTGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1030
 Db 961 ATGTGGAGAGGGGGCTGGCTCTGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTC 1020
 Qy 1031 GGCAACCAACCATCTGTGAGAGGACGCTCATTAACGCCCAAGTGGTCTCACTGCGGCC 1090
 Db 1021 GGCAACCAACCATCTGTGAGAGGACGCTCATTAACGCCCAAGTGGTCTCACTGCGGCC 1080
 Qy 1091 CACTGCTTCTTCTGTGACCCGGGAGAGAGTCTGTGAGGGCTGTGAGAGTGTACGCGGCAAC 1150
 Db 1081 CACTGCTTCTTCTGTGACCCGGGAGAGAGTCTGTGAGGGCTGTGAGAGTGTACGCGGCAAC 1140
 Qy 1151 AGCAACCTGCAACGCTGCTGAGAGGACCTTCCATTTGCCAGATCATCATCAAGCAAT 1210
 Db 1141 AGCAACCTGCAACGCTGCTGAGAGGACCTTCCATTTGCCAGATCATCATCAAGCAAT 1200
 Qy 1211 TACACCGATGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGAC 1270
 Db 1201 TACACCGATGAGAGAGCACTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGAC 1260
 Qy 1271 CTGTCCGCTCATCATCAACCTCTGTGCTTCCCTCCCATGACATGAGACACCTTACCTCAAT 1330
 Db 1261 CTGTCCGCTCATCATCAACCTCTGTGCTTCCCTCCCATGACATGAGACACCTTACCTCAAT 1320
 Qy 1331 GAGACTGCTGTGATCAACGCTTGTGGCAACGAGGAGACATGACATGACATATCCCC 1390
 Db 1321 GAGACTGCTGTGATCAACGCTTGTGGCAACGAGGAGACATGACATGACATATCCCC 1380
 Qy 1391 TTCTCCGGGGGGGGGAGTCAATCTCATGCACTTCAAGAAATGACATGACTTACCTGGTC 1450
 Db 1381 TTCTCCGGGGGGGGGAGTCAATCTCATGCACTTCAAGAAATGACATGACTTACCTGGTC 1440
 Qy 1451 TATGACATTAACCTTACCCCAAGATGATGTCTGTGGGACCTTGTGGGGGAGAGAC 1510
 Db 1441 TATGACATTAACCTTACCCCAAGATGATGTCTGTGGGACCTTGTGGGGGAGAGAC 1500
 Qy 1511 TCCTGCGAGGAGAGACAGCGGGGGGCTTGTGTGTGAGCAAGAAACACGCTGTACTG 1570
 Db 1501 TCCTGCGAGGAGAGACAGCGGGGGGCTTGTGTGTGAGCAAGAAACACGCTGTACTG 1560
 Qy 1571 GCAGGTCTCACCACTGGGGGACAGCGCTGTGGCAAGAAACAAACCTGTGTATCAAC 1630
 Db 1561 GCAGGTCTCACCACTGGGGGACAGCGCTGTGGCAAGAAACAAACCTGTGTATCAAC 1620
 Qy 1631 AAAGTGAAGAGTCTTCTCTGTGATTTACAGAAATGAGAGAGGTCGATTCAGA 1690
 Db 1621 AAAGTGAAGAGTCTTCTCTGTGATTTACAGAAATGAGAGAGGTCGATTCAGA 1680
 Qy 1691 AAATCTTAA 1699
 Db 1681 AAATCTTAA 1689
 RESULT 5
 ADI10392 ID ADI10392 standard; DNA; 1689 BP.
 XX AC ADI10392;
 XX 22-APR-2004 (first entry)
 XX Human cell surface protease coding sequence #12.
 XX therapeutic agent; plasmin; protease specific antigen; PSA;
 KW cell-surface protease-associated disease; cancer; ocular disease;
 KW cardiovascular disease; chronic inflammatory disease; wound;
 KW circulatory disorder; dermatological disorder; rheumatoid arthritis;
 KW psoriasis; diabetic retinopathy; pterygium;
 KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; crest syndrome; solid neoplasm; vascular tumour;
 KW melanoma; Kaposi's sarcoma; human; cell surface protease; gene; ds.

XX OS Homo sapiens.
 XX PN W0200295007-A2.
 XX PD 28-NOV-2002.
 XX PF 23-MAY-2002; 2002W0-US016819.
 XX PR 23-MAY-2001; 2001US-0293267P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Madison EL, Sempke JE, Vlaeuk GP, Kemp SJ, Komandla M, Siev DV;
 XX WP1; 2003-221280/21.
 XX DR P-PSDB; ADI10393.
 XX PT Novel conjugate useful for treating cell-surface protease-associated
 PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
 PT substrate linked to it optionally by a peptidic linker.
 XX PS Claim 9; SEQ ID NO 23; 581bp; English.
 CC The invention comprises a conjugate that consists of a therapeutic agent
 CC and a peptide substrate (optionally linked via linker). The peptide
 CC substrate is proteolytically cleaved by a cell surface protease pr a
 CC soluble, released or shed form of it, to liberate the therapeutic agent,
 CC the conjugate of the invention is not substantially cleaved by plasmin or
 CC protease specific antigen (PSA). The conjugate of the invention is useful
 CC for treating a cell-surface protease-associated disease such as: cancer,
 CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
 CC wounds, circulatory disorders, dermatological disorders, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
 CC scarring from excimer laser surgery, scarring from glaucoma filtering
 CC surgery, macular degeneration, crest syndromes, solid neoplasms, vascular
 CC tumours, melanoma and Kaposi's sarcoma. The present DNA sequence encodes
 CC a human cell surface protease.
 XX SQ Sequence 1689 BP; 367 A; 529 C; 470 G; 323 T; 0 U; 0 Other;
 Query Match 96.5%; Score 1687.4; DB 10; Length 1689;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 11 ATGAGAGGAGCAAGCCAGGGAATGATCTTCAGCAAGAAACACCTTACGCTGAGCATCT 70
 Db 1 ATGAGAGGAGCAAGCCAGGGAATGATCTTCAGCAAGAAACACCTTACGCTGAGCATCT 60
 Qy 71 CCAGCCAGGCAATCTCCAGCTGGGAGACCTCCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 130
 Db 61 CCAGCCAGGCAATCTCCAGCTGGGAGACCTCCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 120
 Qy 131 CCAGCCAGGCAATCTCCAGCTGGGAGACCTCCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 190
 Db 121 CCAGCCAGGCAATCTCCAGCTGGGAGACCTCCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 180
 Qy 181 CCAGCTGTATACCTCCAGGCGGGGCAATCTTCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 240
 Db 191 CCAGCTGTATACCTCCAGGCGGGGCAATCTTCAGGGCCGGGCAATCTCCAGCCAGGCAATCT 250
 Qy 251 CCAGCCCGGGGCAATCTCCGGGCTGTGGCATCACTTTCCAGAGTCTCATCCGAGAGTCA 310
 Db 241 CCAGCCCGGGGCAATCTCCGGGCTGTGGCATCACTTTCCAGAGTCTCATCCGAGAGTCA 300
 Qy 311 TCCGCAAGTCAAGCTGTGAGCAACCTCCCAACAGAGTATCTTGTAGACA 370
 Db 301 TCCGCAAGTCAAGCTGTGAGCAACCTCCCAACAGAGTATCTTGTAGACA 360
 Qy 371 CCAATGGGGGCTGTATCCATTCATATCTCTGCAAGTCAAGCAACAGAGGCC 430
 Db 361 CCAATGGGGGCTGTATCCATTCATATCTCTGCAAGTCAAGCAACAGAGGCC 420

QY	431	ACCAAGGAGAGACCACAGGTACGAGGCTCCACAGTTACCTCGCGGAGGCGCAGAACAG	430
Db	421	ACCAAGGAGAGACCACAGGTACGAGGCTCCACAGTTACCTCGCGGAGGCGCAGAACAG	480
QY	491	CTACCGGCTCATCGGGTGCAGTCTCTCTCTCATTTGCTCGGTGGTTTGCTCATCATCTCT	550
Db	481	CTACCGGCTCATCGGGTGCAGTCTCTCTCTCATTTGCTCGGTGGTTTGCTCATCATCTCT	540
QY	551	TTCCAGTTCTGGCAGGGCCACAGGGATCAGTTACAGAGACAGAGGAGAGCTGTCC	610
Db	541	TTCCAGTTCTGGCAGGGCCACAGGGATCAGTTACAGAGAGCAGAGGAGAGCTGTCCC	600
QY	611	AAGAAGCTGTTCGCTGTGACGGGGTGGTGGACCTGCAAGCTGAAGAGTGAACGAGCTGGC	670
Db	601	AAGAAGCTGTTCGCTGTGACGGGGTGGTGGACCTGCAAGCTGAAGAGTGAACGAGCTGGC	660
QY	671	TGCGTGAAGTTTGACTGGGACAGAGTCTCTGCTTAAATCTACTCTGGGTCTCTCCATCAG	730
Db	661	TGCGTGAAGTTTGACTGGGACAGAGTCTCTGCTTAAATCTACTCTGGGTCTCTCCATCAG	720
QY	731	TGGCTTCCCATCTGTAGCAGCAACTGGAAATGACTCTTACTCAGAGAAACCTGCGACAG	790
Db	721	TGGCTTCCCATCTGTAGCAGCAACTGGAAATGACTCTTACTCAGAGAAACCTGCGACAG	780
QY	791	CTGGGTTTCGAGAGTGTCTACCCGACACCGAGGTTGCCACAGGATTTTGGCAACAGC	850
Db	781	CTGGGTTTCGAGAGTGTCTACCCGACACCGAGGTTGCCACAGGATTTTGGCAACAGC	840
QY	851	TTCTCAATCTTGAGATACCACTCCACCATCCAGGAAAGGCTCCACAGGTCTGAATGGCCT	910
Db	841	TTCTCAATCTTGAGATACCACTCCACCATCCAGGAAAGGCTCCACAGGTCTGAATGGCCT	900
QY	911	TCCGAGCGGTATATCTCCCTCAGTGTTCCCACTGCGGACTGAAGGCGCATGACCGGCGG	970
Db	901	TCCGAGCGGTATATCTCCCTCAGTGTTCCCACTGCGGACTGAAGGCGCATGACCGGCGG	960
QY	971	ATCGTGGAGAGGGGCGCTGAGCTCCGATAGCAAGTGGCTTGGCAATGAGTCTGCACTTC	1030
Db	961	ATCGTGGAGAGGGGCGCTGAGCTCCGATAGCAAGTGGCTTGGCAATGAGTCTGCACTTC	1020
QY	1031	GGCACCAACCCACATCTGTGAGAGGACGCGTCAATTAACGCCAGTGGGAGTCACTGCGCGC	1090
Db	1021	GGCACCAACCCACATCTGTGAGAGGACGCGTCAATTAACGCCAGTGGGAGTCACTGCGCGC	1080
QY	1091	CAGTGCCTCTTCGTGACCCCGGAGAGAGGTCCTGGAAGGCTGGAAGGTGTACGCGGAGCC	1150
Db	1081	CAGTGCCTCTTCGTGACCCCGGAGAGAGGTCCTGGAAGGCTGGAAGGTGTACGCGGAGCC	1140
QY	1151	AGCAACTGTGACCAAGTTGCTGAGGCAAGCTTCATTTGCCGAGATCATACACGAAT	1210
Db	1141	AGCAACTGTGACCAAGTTGCTGAGGCAAGCTTCATTTGCCGAGATCATACACGAAT	1200
QY	1211	TACACCGATGAGGAGAGACGACTATGACATGAGCCGCAATGGCGCTGTCCAGGCCCTAGCC	1270
Db	1201	TACACCGATGAGGAGAGACGACTATGACATGAGCCGCTCATGCGGCTGTCCAGGCCCTAGCC	1260
QY	1271	CTGTCCGCTCATCTCAACCTGTCTGTCTGCTCCCATGCAATGACAGACCTTTAGCCTCAAT	1330
Db	1261	CTGTCCGCTCATCTCAACCTGTCTGTCTGCTCCCATGCAATGACAGACCTTTAGCCTCAAT	1320
QY	1331	GAGACCTGTCTGATATCACAGGCTTTTGGCAAGACCAAGGAGACAGATGACAGACATCCCC	1390
Db	1321	GAGACCTGTCTGATATCACAGGCTTTTGGCAAGACCAAGGAGACAGATGACAGACATCCCC	1380
QY	1391	TTCTCCCGGAGGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCATATGACTATCTGTGTC	1450
Db	1381	TTCTCCCGGAGGAGTGAAGTCAATCTCATGCACTTCAAGAAATGCATATGACTATCTGTGTC	1440
QY	1451	TATGACAGTTACTTTACCCCAAGATGATGTGTCTGGGAGCTTTCGTGGGGGACAGAGAC	1510
Db	1441	TATGACAGTTACTTTACCCCAAGATGATGTGTCTGGGAGCTTTCGTGGGGGACAGAGAC	1500
QY	1511	TCCTGCGAGGAGACAGCGGGGGGCTCTTGTGTGTGAGCAAGAACCGCTGTACTGTG	1570

Db	1501	TCCTGCCAGGAGACACGGGGGGCCCTCTTGTCTGTAGCAGAAACAACCGCTGGTACTCTG	1560
Qy	1571	GCAGGTGTCAACCACTTGGGGGACACAGGCTGTGGCCAGAGAAACAACCTGGTGTATACCC	1630
Db	1561	GCAGGTGTCAACCACTGGGGGACACAGGCTGTGGCCAGAGAAACAACCTGGTGTATACCC	1620
Qy	1611	AAAGTGCACAGAAAGTTCTTCTCTGGATTATTACACGAAGATGCGAGGACAGGTGCCATTACGA	1690
Db	1621	AAAGTGCACAGAAAGTTCTTCTCTGGATTATTACAGCAAGATGCGAGGACAGGTGCCATTACGA	1680
Qy	1691	AAATCCTAA 1699	
Db	1681	AAATCCTAA 1689	
RESULT 6			
ID	ADJ46916	standard; cDNA; 1689 BP.	
AC	ADJ46916;		
DT	06-MAY-2004	(first entry)	
DE	Human transmembrane serine protease (MSP)-related cDNA #2.		
XX	Human; transmembrane serine protease; MTSP; gene; sr;		
KM	cell surface protease; plasmin; prostate specific antigen; PSA;		
KM	proliferative disease; cell-surface protease-associated disease;		
KM	autoimmune disease; inflammatory disease; infectious disease;		
KM	endocrine disease; cancer; ocular disorder; cardiovascular disorder;		
KM	chronic inflammatory disease; wound; circulatory disorder;		
KM	dermatological disorder; restenosis; rheumatoid arthritis; psoriasis;		
KM	diabetic retinopathy; laser surgery scarring;		
KM	glaucoma; filtering surgery infection; macular degeneration;		
KM	CRST syndrome; bacterial infection; viral disease; solid neoplasm;		
KM	vascular tumour; lung; colon; prostate; melanoma; Kaposi's sarcoma.		
XX	Homo sapiens.		
OS	US2004001801-A1.		
PN	01-JAN-2004.		
XX	23-MAY-2002; 2002US-00156214.		
PF	23-MAY-2002; 2002US-00156214.		
XX	(CORV-) CORVAS INT INC.		
PA	Madison ED, Semple JE, Vlausk GP, Kemp SJ, Komandla M, Slav DV;		
PI	WPI; 2004-190126/18.		
XX	P-P8DB; ADJ46917.		
DR	Conjugate useful for treating e.g. cancer, cell-surface protease-		
PT	associated diseases, comprising a peptidic substrate or nucleic acid		
PT	substrate linked to a therapeutic agent through a linker.		
XX	Claim 9; SEQ ID NO 23; 361bp; English.		
PS	The invention relates to a conjugate comprising a therapeutic agent and a		
XX	peptidic substrate or nucleic acid substrate linked to the agent		
CC	optionally through a linker or peptidic linker, where the peptidic		
CC	substrate is proteolytically cleaved by a cell surface protease or a		
CC	soluble, released or shed form conjugate to liberate the agent and the		
CC	conjugate is not substantially cleaved by plasmin or prostate specific		
CC	antigen (PSA). The conjugate is useful for treating a disease, which		
CC	involves administering a conjugate to a subject, where the disease is		
CC	preferably a proliferative diseases or a cell-surface protease-associated		
CC	disease. The diseases include autoimmune diseases, inflammatory diseases,		
CC	infectious diseases and endocrine diseases. The conjugate is useful for		
CC	treating a cell-surface protease-associated disease, which involves		

Query Match	Beat Local Similarity	96.5%	Score 1687.4	DB 14	Length 1689
Matches 1688	Conservative	0	Mismatches	1	Indels 0
					Gaps 0
11	ATGGAGAGGAGCAGGACGACGGGAAATGATCTCCAGCAAGAAACCTTCAGCTGAGGAGCATCT	70			
1	ATGGAGAGGAGCAGGACGACGGGAAATGATCTCCAGCAAGAAACCTTCAGCTGAGGAGCATCT	60			
71	CCAGGCCAGGACATCTCCAGCTGAGGACACCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	130			
61	CCAGGCCAGGACATCTCCAGCTGAGGACACCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	120			
131	CCAGGCCAGGACATCTCCAGCTGAGGACACCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	190			
121	CCAGGCCAGGACATCTCCAGCTGAGGACACCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	180			
191	CCAGCTGATGACATCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	250			
181	CCAGCTGATGACATCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	240			
251	CCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	310			
241	CCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGGCCAGGACATCT	300			

QY	311	TTCCGCAGGTCAGCTTCGGGTGACAACTCCCAACCAAGATGTACCTTTGTAGGCAACA	370
Db	301	TTCCGCAGGTCAGCTTCGGGTGACAACTCCCAACCAAGATGTACCTTTGTAGGCAACA	360
QY	371	CCAAGTGGGGGCTGTATCCCATTCGGATCATCTCTCCGACGATCAGCACCAAGCAACCGAGGCC	430
Db	361	CCAAGTGGGGGCTGTATCCCATTCGGATCATCTCTCCGACGATCAGCACCAAGCAACCGAGGCC	420
QY	431	ACCAAGGAGAGCCCAAGGTATCGAGCTCTGCCAATTCACCTCTGGCGGGAGGGCCAGAACACAG	490
Db	421	ACCAAGGAGAGCCCAAGGTATCGAGCTCTGCCAAGTTCACCTCTGGCGGGAGGGCCAGAACACAG	480
QY	491	CTACCCGCTCATTCGGGGTACGTGCTCTCTCTCTCATTTGCGCTGTGGTTTGTGCTCATCATCTCTC	550
Db	481	CTACCCGCTCATTCGGGGTACGTGCTCTCTCTCTCTCATTTGCGCTGTGGTTTGTGCTCATCATCTCTC	540
QY	551	TTCCAGATTCTGGCAGGGGCCACACAGGGATCAGGTATCAAGAGCAGAGGAGAGCTGTCCC	610
Db	541	TTCCAGATTCTGGCAGGGGCCACACAGGGATCAGGTATCAAGAGCAGAGGAGAGCTGTCCC	600
QY	611	AAGAACCTCTGTGCGCTGTGACGGGGTGGTGGACCTGCAGAGCGTGAAGATGTACGAGCTGGGC	670
Db	601	AAGAACCTCTGTGCGCTGTGACGGGGTGGTGGACCTGCAGAGCTGAAGATGTACGAGCTGGGC	660
QY	671	TGCGTGAAGTTTGACTGTGACCAAGTCTCTGCTTTAAATCTACTCTGGGCTCTCCCATCAG	730
Db	661	TGCGTGAAGTTTGACTGTGACCAAGTCTCTGCTTTAAATCTACTCTGGGCTCTCCCATCAG	720
QY	731	TGGCTTTCCCATCTGTATGACAGCACTGGATATGATCTCTCATCAGAGAAAGACTGTGCACAG	790
Db	721	TGGCTTTCCCATCTGTATGACAGCACTGGATATGATCTCTCATCAGAGAAAGACTGTGCACAG	780
QY	791	CTGGGTTTTCAGAGATGTCTCACCCGACCAACGAGGTTTCCCAACGAGGATTTTGTCCAAAGC	850
Db	781	CTGGGTTTTCAGAGATGTCTCACCCGACCAACGAGGTTTCCCAACGAGGATTTTGTCCAAAGC	840
QY	851	TTCTCAATCTTTAGATATCAACTCCACATCCAGAAAGCTTCCACAGGCTTGAAATGCTCCT	910
Db	841	TTCTCAATCTTTAGATATCAACTCCACATCCAGAAAGCTTCCACAGGCTTGAAATGCTCCT	900
QY	911	TCCCAAGGGATATATCTCCCTCCAGTGTGCCACCTGGGGACCTGAGGGGCACTGACCGGGCGG	970
Db	901	TCCCAAGGGATATATCTCCCTCCAGTGTGCCACCTGGGGGCACTGAGGGGCACTGACCGGGCGG	960
QY	971	ATCGTGGAGGGGGCGCTGAGCCTCGGATATGCAAGTGAGCTTGTGCAAGTGTCTGCACTTC	1030
Db	961	ATCGTGGAGGGGGCGCTGAGCCTCGGATATGCAAGTGTGCTTGTGCAAGTGTCTGCACTTC	1020
QY	1031	GGCAGCCACCATCTGTGTGAGGCAAGCTCATTTGACGCCACATGTGGGTGTCTCATCTGGCGCC	1090
Db	1021	GGCAGCCACCATCTGTGTGAGGCAAGCTCATTTGACGCCACATGTGGGTGTCTCATCTGGCGCC	1080
QY	1091	CACATGCTTTCTTGGTGACCCGGGGAGAAAGTCTGTGAGGGGCTGTGAAGGTGTACGCGGGCAC	1150
Db	1081	CACATGCTTTCTTGGTGACCCGGGGAGAAAGTCTGTGAGGGGCTGTGAAGGTGTACGCGGGCAC	1140
QY	1151	AGCAACCTGCACCAAGTTGCTTGAGGACAGCTTCCATTGCGAGATCATCATCAGCAAT	1210
Db	1141	AGCAACCTGCACCAAGTTGCTTGAGGACAGCTTCCATTGCGAGATCATCATCAGCAAT	1200
QY	1211	TACACCGATAGAGAGGAGCGACTATATGATATGGCCCTCATATGGGGCTGTCCAAAGCCCTTAC	1270
Db	1201	TACACCGATAGAGAGGAGCGACTATATGATATGGCCCTCATATGGGGCTGTCCAAAGCCCTTAC	1260
QY	1271	CTGTCCGCTCATATCACCCTGTGCTTGCCTCCCATATGATGAGACAGACTTTTAACTCAAT	1330
Db	1261	CTGTCCGCTCATATCACCCTGTGCTTGCCTCCCATATGATGAGACAGACTTTTAACTCAAT	1320
QY	1331	GAGACCTGTGATATCAGAGCTTTGGCAAGACCAAGGAGACAGATGACAAAGACTTCCCC	1390
Db	1321	GAGACCTGTGATATCAGAGCTTTGGCAAGACCAAGGAGACAGATGACAAAGACTTCCCC	1380

QY 1391 TTCCTCCGGAGGAGTGCAGTCAATCTCATCTTCAAGAAATGCAATGACTTCTGATC 1450
XX
Db 1381 TTCTTCGGGGAGTGCGAGTCAATCTCATCTTCAAGAAATGCAATGACTTCTGATC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGGGGGAGAGAC 1510
1441 TATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGGGGGAGAGAC 1500
QY 1511 TCTCGCCAGGAGACAGCGGGGGCCCTTCTGTGTGAGCAACCAACCTGTTACCTG 1570
1501 TCTCGCCAGGAGACAGCGGGGGCCCTTCTGTGTGAGCAACCAACCTGTTACCTG 1560
QY 1571 GCAGGTTCACACAGCTGGGGGACAGGCTGTGGCCAGAAACCAACCTGTTACAC 1630
1561 GCAGGTTCACACAGCTGGGGGACAGGCTGTGGCCAGAAACCAACCTGTTACAC 1620
QY 1631 AAAGTGAAGAGTTCTTCTCTGATTTACAGCAAGATGAGAGCGAGTTCATCAGA 1690
1621 AAAGTGAAGAGTTCTTCTCTGATTTACAGCAAGATGAGAGCGAGTTCATCATA 1680
QY 1691 AAATCTTAA 1699
1681 AAATCTTAA 1689

RESULT 8
AD275551
ID AD275551 standard; cDNA; 2380 BP.

XX AD275551;

DT 14-JUL-2005 (first entry)

XX Human mosaic serine protease (MSP) cDNA.

XX
XX cardiant; endocrine-gen.; metabolic; gastrointestinal; hemostatic;
XX respiratory-gen.; nootropic; neuroprotective; uropathic; gynecological;
XX antiinflammatory; cytostatic; antianemic; diagnostic; pharmaceutical;
XX cardiovascular disease; cardiovascular-gen.; endocrine-gen.; metabolic;
XX metabolic disorder; hematological disease; respiratory disease;
XX respiratory-gen.; neurological disease; andrology; endocrine disease;
XX gastrointestinal disease; gynecology and obstetrics; inflammation;
XX cancer; cytostatic; neoplasm; mosaic serine protease; MSP; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT CDS 88..1833

XX FT /product /tag=a

XX FT /transl_except= (pos:862..864, aa:Arg)

XX PN W02005040401-A2.

XX 06-MAY-2005.

XX 02-OCT-2004; 2004MO-EP011015.

XX 17-OCT-2003; 2003EP-00023809.

XX (FARB) BAYER HEALTHCARE AG.

XX Golz S, Brueggemeier U, Geerts A;

XX MPI; 2005-355849/36.

XX P-PSDB; AD275552.

XX Screening for therapeutic agents for treating cardiovascular, metabolic,
XX hematological, neurological, urological or respiratory diseases, or
XX inflammation by contacting a test compound with a mosaic serine protease
XX polypeptide.

PS Disclosure; SEQ ID NO 1; 98pp; English.

XX The invention describes a method of screening for therapeutic agents for
XX treating cardiovascular, endocrinological, metabolic, gastrointestinal,
XX hematological, respiratory, neurological, urological and reproduction
XX disorders, inflammation, or cancer in a mammal comprising contacting a
XX test compound with a mosaic serine protease (MSP) polypeptide. The method
XX comprises: contacting a test compound with a MSP polypeptide; and
XX detecting binding of the test compound to the MSP polypeptide. Also
XX described are: a method of diagnosing any of the diseases cited above in
XX a mammal comprising determining the amount of a MSP polynucleotide in a
XX sample taken from the mammal; and determining the amount of MSP
XX polynucleotide in healthy and/or diseased mammals; a pharmaceutical
XX composition for the treatment of any of the diseases cited above,
XX comprising a therapeutic agent which binds to, or regulates the activity
XX of, a MSP polypeptide; or a MSP polynucleotide or polypeptide; a method
XX for preparing the pharmaceutical composition comprising identifying a
XX regulator of MSP; determining whether the regulator ameliorates the
XX symptoms the disease in a mammal; and combining of the regulator with an
XX acceptable pharmaceutical carrier; and a regulator of MSP for the
XX regulation of MSP activity in a mammal having any of the diseases cited
XX above. The method is useful for screening for therapeutic agents for
XX treating cardiovascular, endocrinological, metabolic, gastrointestinal,
XX hematological, respiratory, neurological, urological and reproduction
XX disorders, inflammation, or cancer in a mammal. The compositions are
XX useful for treating such diseases. This sequence encodes human mosaic
XX serine protease (MSP).

XX Sequence 2380 BP; 540 A; 692 C; 627 G; 521 T; 0 U; 0 Other;

XX Query Match 95.6%; Score 1670.4; DB 14; Length 2380;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCAGAGACCATGAGAGGAGACCCAGCGAATGATCTCCAGCAAGAACCTTCAGC 60
Db 78 CTCAGAGACCATGAGAGGAGAGACCCAGCGAATGATCTCCAGCAAGAACCTTCAGC 137
QY 61 TGGAGCATCTCCAGCCGAGCATCTCCAGTGGAGACCTTCAGCCGGGATCTCCAGC 120
138 TGGAGCATCTCCAGCCGAGCATCTCCAGTGGAGACCTTCAGCCGGGATCTCCAGC 197
QY 121 CCAAGCATCTCCAGCCGAGCATCTCCAGTGGAGACCTTCAGCCGGGATCTCCAGC 180
198 CCAAGCATCTCCAGCCGAGCATCTCCAGTGGAGACCTTCAGCCGGGATCTCCAGC 257
QY 181 CCAAGCATCTCCAGCTGTACACCTCCAGCCGAGCATCTCCAGCCGGGATCTCCAGC 240
258 CCAAGCATCTCCAGCTGTACACCTCCAGCCGAGCATCTCCAGCCGGGATCTCCAGC 317
QY 241 CCAAGCATCTCCAGCCGAGCATCTCCAGCTGTACACCTTCAGCTTCATCTCCAG 300
318 CCAAGCATCTCCAGCCGAGCATCTCCAGCTGTACACCTTCAGCTTCATCTCCAG 377
QY 301 CAGGTCAATCCGCGAGCTGAGCTGAGTGAACCTCCCAACGAGAGTACCTTGT 360
378 CAGGTCAATCCGCGAGCTGAGCTGAGTGAACCTCCCAACGAGAGTACCTTGT 437
QY 361 TAGAGCAACACAGTGGGGCTGTACCATCCATCTCTGTCAGAGTCCAGCAGC 420
438 TAGAGCAACACAGTGGGGCTGTACCATCCATCTCTGTCAGAGTCCAGCAGC 497
QY 421 AACCAAGGCCACAGGAGAGCCAGGATGAGCTGAGTCCCAAGTTCACTGCGGAGGG 480
498 AACCAAGGCCACAGGAGAGCCAGGATGAGCTGAGTCCCAAGTTCACTGCGGAGGG 557
QY 481 CCAAGAGCAGTACCGCTCATCGGGTGGCTCTCTCATTTGCGTGGTGGTGGCT 540
558 CCAAGAGCAGTACCGCTCATCGGGTGGCTCTCTCATTTGCGTGGTGGTGGCT 617
QY 541 CATCATCTCTTCCAGTTCTGGCAGGGCCACAGAGGATCAGGTACAAGAGCAGAGGGA 600
618 CATCATCTCTTCCAGTTCTGGCAGGGCCACAGAGGATCAGGTACAAGAGCAGAGGGA 677

QY 601 GAGCTGTCCCAAGCAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 660
 DB 678 GAGCTGTCCCAAGCAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 737
 QY 661 CGAGCTGGGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 720
 DB 738 CGAGCTGGGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 797
 QY 721 CTCCATCAGTGGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 780
 DB 798 CTCCATCAGTGGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 857
 QY 781 CTGCGACAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 840
 DB 858 CTGCGACAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 917
 QY 841 TGGCAACAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 900
 DB 918 TGGCAACAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 977
 QY 901 TGAATGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 960
 DB 978 TGAATGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1037
 QY 961 GACCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1020
 DB 1038 GACCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1097
 QY 1021 TCGCACTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1080
 DB 1098 TCGCACTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1157
 QY 1081 CACTGCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1140
 DB 1158 CACTGCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1217
 QY 1141 CGCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1200
 DB 1218 CGCGGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1277
 QY 1201 CAACAGCACTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1260
 DB 1278 CAACAGCACTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1337
 QY 1261 GCGGCTGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1320
 DB 1338 GCGGCTGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1397
 QY 1321 TAGCTCAATGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1380
 DB 1398 TAGCTCAATGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1457
 QY 1381 GACATCCCTCTCTGCGGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1440
 DB 1458 GACATCCCTCTCTGCGGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1517
 QY 1441 CTACTGTGATGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1500
 DB 1518 CTACTGTGATGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1577
 QY 1501 GCGGAGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1560
 DB 1578 GCGGAGAGCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1637
 QY 1561 CTGCTTACCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1620
 DB 1638 CTGCTTACCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1697
 QY 1621 TGTGTACACCAAGTGAAGATGTTCCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1672
 DB 1698 TGTGTACACCAAGTGAAGATGTTCCTGTGAGGTGTGATCTGCAAGCTGTTCCTGTGACGAGGTGTGTGATCTGCAAGCTGAAGATGA 1749

RESULT 9
 ADE31742
 ID ADE31742 standard; DNA; 2393 BP.
 XX
 AC ADE31742;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human 2914 gene #SEQ ID 99.
 XX
 KM Antihistaminergic; carbanil; vasodilator; anti-inflammatory;
 KM chromolytic; antihistaminic; antianxiety; hypotensive; gene therapy;
 KM cardiovascular; disorder; ischaemia; aortic bending;
 KM vascular heart disease; endocarditis; atrial fibrillation; heart failure;
 KM angina; cardiomyopathy; cardiac death; gene; de.
 XX
 OS Homo sapiens.
 XX
 PN W02003065984-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 29-JAN-2003; 2003MO-US002571.
 XX
 PR 01-FEB-2002; 2002US-0353224P.
 PR 15-MAR-2002; 2002US-0364529P.
 PR 19-APR-2002; 2002US-0373611P.
 PR 29-APR-2002; 2002US-0376287P.
 PR 12-JUN-2002; 2002US-0388080P.
 PR 24-JUN-2002; 2002US-0390971P.
 PR 03-JUL-2002; 2002US-0394130P.
 PR 10-JUL-2002; 2002US-0394797P.
 PR 21-AUG-2002; 2002US-0404904P.
 PR 23-AUG-2002; 2002US-0405450P.
 PR 04-SEP-2002; 2002US-0408070P.
 PR 06-NOV-2002; 2002US-0424300P.
 PR 05-DEC-2002; 2002US-0431042P.
 PR 05-DEC-2002; 2002US-0431079P.
 PA
 (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Logan TU, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
 PI Stegiano N, Perodin J, Rodrigue-Way A;
 XX
 DR MPI; 2003-731468/69.
 DR P-PSDB; ADE31743.
 XX
 PT Identifying a compound capable of treating a cardiovascular disorder
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193
 PT polypeptide or nucleic acid.
 XX
 PS Disclosure; SEQ ID NO 99; 328bp; English.
 XX
 CC The invention relates to a method for identifying a compound capable of
 CC treating a cardiovascular disorder. The present invention identifies the
 CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
 CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
 CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
 CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
 CC 2888, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
 CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
 CC 6585 genes in cardiovascular disease states. The methods are useful in
 CC diagnosing, preventing and treating cardiovascular disorders, such as
 CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
 CC revascularization, arterial inflammation, vascular wall remodeling, coronary
 CC microembolism, tachycardia, bradycardia, pressure overload, aortic
 CC bending, coronary artery ligation, vascular heart disease, valvular
 CC disease, including but not limited to, valvular degeneration caused by
 CC calcification, rheumatic heart disease, endocarditis, or complications of
 CC artificial valves, atrial fibrillation, long-QT syndrome, congestive
 CC heart failure, sinus node dysfunction, angina, heart failure,

CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
 CC including but not limited to, pericardial effusion and pericarditis;
 CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
 CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
 CC artery spasm, ischemic disease, arrhythmia, sudden cardiac death and
 CC cardiovascular developmental disorders. The methods may also be used for
 CC identifying compounds that modulate cardiovascular disorders. Sequences
 CC given in ADE31644-ADE31769 represent the genes and proteins that may be
 CC regulated by a compound of the invention.

XX Sequence 2393 BP, 553 A, 692 C, 627 G, 521 T, 0 U, 0 Other;

Query Match 95.6%; Score 1670.4; DB 10; Length 2393;

Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCAGAGACATGAGAGAGGAGACGACGAGAAATGATCTCCAGACAGAACACCTTCAGC 60
 DB 78 CTCAGAGACATGAGAGAGGAGACGACGAGAAATGATCTCCAGACAGAACACCTTCAGC 137
 QY 61 TGGAGCATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 120
 DB 138 TGGAGCATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 197
 QY 121 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 180
 DB 198 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 257
 QY 181 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 240
 DB 258 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 317
 QY 241 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 300
 DB 318 CCAGGACATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 377
 QY 301 CAGGTCAATCATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 360
 DB 378 CAGGTCAATCATCTCCAGACCCAGAGCATCTCCAGCTGGGACACCTCCAGGCCGGGATCTCCAGC 437
 QY 361 TAGAGCAACACAGTGGGGGCTGTACCCATCCATCTCTCCAGTCCAGGATCAGACCCAGC 420
 DB 438 TAGAGCAACACAGTGGGGGCTGTACCCATCCATCTCTCCAGTCCAGGATCAGACCCAGC 497
 QY 421 AACCAAGGACCCAGAGGAGAGCCAGAGTACAGAGCTGCCCAGATTCACCTGGGGAGAGG 480
 DB 498 AACCAAGGACCCAGAGGAGAGCCAGAGTACAGAGCTGCCCAGATTCACCTGGGGAGAGG 557
 QY 481 CCAGAAGAGTACCGCTCATCGGGTGGGTCTCTCTCATTTGCCCCGTGGTGGTTTGGCT 540
 DB 558 CCAGAAGAGTACCGCTCATCGGGTGGGTCTCTCTCATTTGCCCCGTGGTGGTTTGGCT 617
 QY 541 CATCATCTCTTCCAGATCTGAGAGGAGCCACAGAGGATCAGATTCAGAGAGAGGAG 600
 DB 618 CATCATCTCTTCCAGATCTGAGAGGAGCCACAGAGGATCAGAGAGAGAGAGGAG 677
 QY 601 GAGCTGTCCCAAGACCGTGTCTGTGACGGGGTGGTGGATCTGCAAGCTGGAAGATGA 660
 DB 678 GAGCTGTCCCAAGACCGTGTCTGTGACGGGGTGGTGGATCTGCAAGCTGGAAGATGA 737
 QY 661 CGAGCTGGGGCTGGGAGATTTGACTGGGACAAAGTCTTGCTTAAATCTACTCTGGGCTC 720
 DB 738 CGAGCTGGGGCTGGGAGATTTGACTGGGACAAAGTCTTGCTTAAATCTACTCTGGGCTC 797
 QY 721 CTCCCATAGTGGCTTCCCATCTGTAGAGCACTGGAATGACTCTTCTCAGAGAAAGAC 780
 DB 798 CTCCCATAGTGGCTTCCCATCTGTAGAGCACTGGAATGACTCTTCTCAGAGAAAGAC 857
 QY 781 CTGCAAGAGCTGGGTTTCCAGAGAGCTCCAGAGCAACGAGAGTTGGCCACAGGAGATT 840
 DB 858 CTGCAAGAGCTGGGTTTCCAGAGAGCTCCAGAGCAACGAGAGTTGGCCACAGGAGATT 917
 QY 841 TGCCAAACAGCTTCTCAATCTTGAGATACACTCCACATCCAGAGAAAGCTTCCACAGGCTC 900

DB 918 TGCCAAACAGCTTCTCAATCTTGAGATACACTCCACATCCAGAGAAAGCTTCCACAGGCTC 977
 QY 901 TGAATGCCCTTCCAGAGGATATATCTCCCTCCAGATTTCCCATCTGCCGACTGAGAGGCAT 960
 DB 978 TGAATGCCCTTCCAGAGGATATATCTCCCTCCAGATTTCCCATCTGCCGACTGAGAGGCAT 1037
 QY 961 GACCGGCGGATCTGGGAGAGGGGCGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG 1020
 DB 1038 GACCGGCGGATCTGGGAGAGGGGCGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAG 1097
 QY 1021 TTTGCACTTGGGACCAACCAATCTGTGAGAGGACGCTCATTTGACGCCAGTGGGCT 1080
 DB 1098 TTTGCACTTGGGACCAACCAATCTGTGAGAGGACGCTCATTTGACGCCAGTGGGCT 1157
 QY 1081 CACTGCCGCCCACTGCTTTCTTCTGACCCCGGAGAAAGTCTCTGAGAGGCTTGAAGTGA 1140
 DB 1158 CACTGCCGCCCACTGCTTTCTTCTGACCCCGGAGAAAGTCTCTGAGAGGCTTGAAGTGA 1217
 QY 1141 CGCGGGACACGACAACTGACAGCTTGGCTGAGGAGGAGGCTTGGCCGAGATCATAT 1200
 DB 1218 CGCGGGACACGACAACTGACAGCTTGGCTGAGGAGGAGGCTTGGCCGAGATCATAT 1277
 QY 1201 CAACAGCAATTACACCGATGAGAGAGAGAGATGACATGACCTCTCAATGCGGCTGTCCAA 1260
 DB 1278 CAACAGCAATTACACCGATGAGAGAGAGAGATGACATGACCTCTCAATGCGGCTGTCCAA 1337
 QY 1261 GCCCTTGAACCTTGTCCGCTCATATCCACCTGCTTGGCTCCCATGATGACAGACCTT 1320
 DB 1338 GCCCTTGAACCTTGTCCGCTCATATCCACCTGCTTGGCTCCCATGATGACAGACCTT 1397
 QY 1321 TAGCCTCAATGAGACCTGCTGATGATCAGAGGCTTGGCAAGACAGGAGAGATGACAA 1380
 DB 1398 TAGCCTCAATGAGACCTGCTGATGATCAGAGGCTTGGCAAGACAGGAGAGATGACAA 1457
 QY 1381 GACATCCCTCTCTCCGAGAGGTGAGGTCAATCTGATGACATTCAGAAATGCAATGA 1440
 DB 1458 GACATCCCTCTCTCCGAGAGGTGAGGTCAATCTGATGACATTCAGAAATGCAATGA 1517
 QY 1441 CTACTTGGTCTATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGGG 1500
 DB 1518 CTACTTGGTCTATGACAGTTACCTTACCCCAAGATGATGTGCTGGGACCTTCTGGG 1577
 QY 1501 GGGCAGAGACTCCCTCCAGAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACACG 1560
 DB 1578 GGGCAGAGACTCCCTCCAGAGGAGACAGCGGGGGGCTCTTGTGTGAGCAGAAACACG 1637
 QY 1561 CTGTTACTTGGCAGAGTGTCAACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAACTGG 1620
 DB 1638 CTGTTACTTGGCAGAGTGTCAACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAACTGG 1697
 QY 1621 TGTGTACACCAAGTACAGAAAGTTCTTCCCTGGATTTACAGCAAGATGAG 1672
 DB 1698 TGTGTACACCAAGTACAGAAAGTTCTTCCCTGGATTTACAGCAAGATGAG 1749

RESULT 10
 AAD05797
 ID AAD05797 standard; DNA, 2067 BP.
 XX AAD05797;
 XX 31-JUL-2001 (first entry)

Human transmembrane serine protease (Endothelinase 2-L) DNA.
 DE Human, endothelinase 2-L; protease domain; cytosolic; vulnerable; wound;
 KW neotrophic; peridontitis; dermatological disorder; gene therapy; scar;
 KW angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KW chronic inflammatory disease; ocular disorder; circulatory disorder;
 KW chest syndrome; atherosclerosis; haemangiomatosis; diabetes mellitus;
 KW liver cirrhosis; osteoradionecrosis; systemic sclerosis; oesophagel;
 KW inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;

KW systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
XX transmembrane serine protease; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..2067
FT /tag= a
FT /product= "Human endothelase 2-L protein"
PN MO200136604-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000MO-US031803.
XX
XX 18-NOV-1999; 99US-0166391P.
XX 22-SEP-2000; 2000US-0234840P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Ong EO;
XX
XX WPI: 2001-335601/35.
XX P-PSDB; AA001944.
XX
XX
XX New nucleic acid encoding a protein comprising endothelase activity
XX useful in the prevention and treatment of e.g. vascular malformations,
XX cardiovascular disorders, and chronic inflammatory disease.
XX
XX Example 2, Page 139-142; 152pp; English.
XX
XX The present DNA sequence encodes human long form transmembrane serine
XX protease (Endothelase 2-L) protein. The invention relates to an
XX endothelase protein, endothelase protease domain and their
XX corresponding nucleic acid molecules. An endothelase protein or protease
XX domain of it is useful for the treatment and diagnosis of disorders
XX associated with aberrant angiogenesis or undesired neovascularisation.
XX The undesired angiogenesis is associated with disorders selected from
XX solid neoplasm, vascular malformations and cardiovascular disorders such
XX as angiofibroma, angiolipoma, atherosclerosis, restenosis/reperfusion
XX injury, arteriovenous malformations, haemangiomas and vascular
XX adhesions, dyschondroplasia with vascular hamartomas (Palucci's
XX syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
XX syndrome) and Von Hippel Lindau syndrome, chronic inflammatory bowel
XX disease, nonhealing fractures, periodontitis, psoriasis, rheumatoid
XX arthritis, venous stasis ulcers, granulations-burns, hypertrophic scars,
XX liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
XX granuloma and systemic sclerosis and aberrant wound repairs, circulatory
XX disorders Raynaud's phenomenon, crest syndromes such as calcinosis,
XX oesophagial, dyomeolysis, sclerodactyly and teangiectasis, dermatological
XX disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
XX vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Fort-wine
XX stain, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
XX and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
XX caused by ocular neovascular disease, corneal graft neovascularisation,
XX macular degeneration, retinopathy of prematurity, retrolental fibroplasia
XX and corneal neovascularisation. The nucleic acids of the invention are
XX also used in gene therapy. The invention also provides method for
XX screening compounds that modulate angiogenesis
XX
XX Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;
XX
XX Query Match 95.2%; Score 1663.4; DB 4; Length 2067;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB
61 CCAGCCAGGCGATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCGATCT
QY
131 CCAGCCAGGCGATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCGATCT
DB
121 CCAGCCAGGCGATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCGATCT
QY
191 CCAGCTGGTACCTCCAGGCGGGGCACTCTCCAGCCCGGGCATCTCCAGCCAGGCGATCT
DB
181 CCAGCTGGTACCTCCAGGCGGGGCACTCTCCAGCCCGGGCATCTCCAGCCAGGCGATCT
QY
251 CCAGCCCGGGGATCTCCGGGCTTGGGATCACTTTCCAGGTCCTCATCCGGAGGTCATCA
DB
241 CCAGCCCGGGGATCTCCGGGCTTGGGATCACTTTCCAGGTCCTCATCCGGAGGTCATCA
QY
311 TCCGCGAGGTCAGGCTCGGTGACAACTCCGCCAACAGAGTACCTTTGATAGACAA
DB
301 TCCGCGAGGTCAGGCTCGGTGACAACTCCGCCAACAGAGTACCTTTGATAGACAA
QY
371 CCAGTGGGGGCTGTACCATCCGATCATCTCTCCAGGTCAGGACGCAACGAGGGCC
DB
361 CCAGTGGGGGCTGTACCATCCGATCATCTCTCCAGGTCAGGACGCAACGAGGGCC
QY
431 ACCAGGGAGAGCCAGAGTACGAGGCTGCGCCAAAGTTACCTGGCGGGAGGCCAAGACG
DB
421 ACCAGGGAGAGCCAGAGTACGAGGCTGCGCCAAAGTTACCTGGCGGGAGGCCAAGACG
QY
491 CTACCGCTCATCGGGTACGCTGCTCCTCCTCAATGSCCTGGTGGTTCGTCATCATCTC
DB
481 CTACCGCTCATCGGGTACGCTGCTCCTCCTCAATGSCCTGGTGGTTCGTCATCATCTC
QY
551 TTCAGTCTTGGACGAGGCGCACACAGAGATCGATGACAGGACGAGAGAGAGTGTCC
DB
541 TTCAGTCTTGGACGAGGCGCACACAGAGATCGATGACAGGAGAGAGAGTGTCC
QY
611 AAGACGCTGTTGCTGTGACGAGGCTGTGACCTGCAAGCTGAAAGTGAAGTGAAGTGGCC
DB
601 AAGACGCTGTTGCTGTGACGAGGCTGTGACCTGCAAGCTGAAAGTGAAGTGAAGTGGCC
QY
671 TGCCTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT
DB
661 TGCCTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT
QY
731 TGCCTTCCCATCTGTACAGAGCACTGGAATGACTCTACTGAGAGAACTGTGCGACAG
DB
721 TGCCTTCCCATCTGTACAGAGCACTGGAATGACTCTACTGAGAGAACTGTGCGACAG
QY
791 CTGGGTTTCCAGAGTGTCTACCCGGAACAGAGGTTGCCACAGGGAATTTTGCCAAAGC
DB
781 CTGGGTTTCCAGAGTGTCTACCCGGAACAGAGGTTGCCACAGGGAATTTTGCCAAAGC
QY
851 TTTCATCTTGAATGATCAACTCCACCATCCAGAGAAAGCTCCACAGGTCGATGCGCT
DB
841 TTTCATCTTGAATGATCAACTCCACCATCCAGAGAAAGCTCCACAGGTCGATGCGCT
QY
911 TCCAGGCGATATCTCCCTTCACTGTTCCTCACTGCGGACTGAGGCGCATGACCGGGCGG
DB
901 TCCAGGCGATATCTCCCTTCACTGTTCCTCACTGCGGACTGAGGCGCATGACCGGGCGG
QY
971 ATCGTGGAGGGGCGCTGCGCTCGGATGAGCAAGTGGCTTGGCAAGTGAAGTCTGACATT
DB
961 ATCGTGGAGGGGCGCTGCGCTCGGATGAGCAAGTGGCTTGGCAAGTGAAGTCTGACATT
QY
1031 GGCACCAACCCACATCTGTGAGGACGCTCATTTGACGCGCCAGTGGGTCATCTGCGGCC
DB
1021 GGCACCAACCCACATCTGTGAGGACGCTCATTTGACGCGCCAGTGGGTCATCTGCGGCC
QY
1091 CACTGCTTCTTGTGACCGCGGAGAGAGTCTGTGAGGAGTGTGAGAGGAGTGTGAGAGGAC
DB
1081 CACTGCTTCTTGTGACCGCGGAGAGAGTCTGTGAGGAGTGTGAGAGGAGTGTGAGAGGAC
QY
1151 AGCAACTGCAACGTTGCTGTGAGGAGGCTTCCATTGCGAGATCATCATCAAGCAAT

D 1141 AGCAACCTGCACGATTGCTGAGGACGCTCCATGTCGCCGATCATCATCAACAGCAAT 1200
Q 1211 TACACCGATGAGGAGGAGCATATGACATGCGCCCTCATGCGGCTGTCCAGACCCCTTGACC 1270
D 1201 TACACCGATGAGGAGGAGCATATGACATGCGCCCTCATGCGGCTGTCCAGACCCCTTGACC 1260
Q 1271 CTGACCGGTCATCAACCCCTGCTGCTCCCGATGATGAGAGACCTTATGCTCAAT 1330
D 1261 CTGACCGGTCATCAACCCCTGCTGCTCCCGATGATGAGAGACCTTATGCTCAAT 1320
Q 1331 GAGACCTGCTGATCAGAGGCTTTGGCAAGACAGGAGACAGATGACAGACATCCCCC 1390
D 1321 GAGACCTGCTGATCAGAGGCTTTGGCAAGACAGGAGACAGATGACAGACATCCCCC 1380
Q 1391 TTCTCTCCGGGAGGTCAGAGTAACTCATGCACTTCAAGAAATGCAATGATCTTGGTC 1450
D 1381 TTCTCTCCGGGAGGTCAGAGTAACTCATGCACTTCAAGAAATGCAATGATCTTGGTC 1440
Q 1451 TATGACAGTTACCTTACCCCAAGATGATGCTGCTGGGACCTTCGTCGGGGCAGAGAC 1510
D 1441 TATGACAGTTACCTTACCCCAAGATGATGCTGCTGGGACCTTCGTCGGGGCAGAGAC 1500
Q 1511 TCCTGCGACGAGGAGACAGCGGGGCTCTTGTCTGTGAGCAGAAACAACCGCTGTACCTG 1570
D 1501 TCCTGCGACGAGGAGACAGCGGGGCTCTTGTCTGTGAGCAGAAACAACCGCTGTACCTG 1560
Q 1571 GCAGGTCACACAGCTGGGGCAGAGGCTGTGGCCAGAGAAACAACCTGTGTGTACACC 1630
D 1561 GCAGGTCACACAGCTGGGGCAGAGGCTGTGGCCAGAGAAACAACCTGTGTGTACACC 1620
Q 1631 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGC 1675
D 1621 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGC 1665
RESULT 11
AD110394
ID AD110394 standard; DNA; 2067 BP.
AC AD110394;
XX
XX 22-APR-2004 (first entry)
DE Human cell surface protease coding sequence #13.
XX
XX therapeutic agent; plasmin; protease specific antigen; PSA;
KM cell-surface protease-associated disease; cancer; ocular disease;
KM cardiovascular disease; chronic inflammatory disease; wound;
KM circulatory disorder; dermatological disorder; rheumatoid arthritis;
KM psoriasis; diabetic retinopathy; pterygium;
KM excimer laser surgery scarring; glaucoma filtering surgery scarring;
KM macular degeneration; cress syndrome; solid neoplasm; vascular tumor;
KM melanoma; Kaposi's sarcoma; human; cell surface protease; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200295007-A2.
XX
XX 28-NOV-2002.
XX
XX 23-MAY-2002; 2002WO-US016819.
XX
XX 23-MAY-2001; 2001US-0293267P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Sempke JB, Vlausk GP, Kemp SJ, Komandla M, Siev DV;
PI WPI; 2003-221280/21.
XX
XX P-PSDB; AD110395.
XX
XX Novel conjugate useful for treating cell-surface protease-associated
PT disease, comprises a therapeutic agent and a peptidic or nucleic acid

PT substrate linked to it optionally by a peptidic linker.
XX
XX Claim 9; SEQ ID NO 25; 581bp; English.
XX
XX The invention comprises a conjugate that consists of a therapeutic agent
CC and a peptide substrate (optionally linked via linker). The peptide
CC substrate is proteolytically cleaved by a cell surface protease pr a
CC soluble, released or shed form of it, to liberate the therapeutic agent,
CC the conjugate of the invention is not substantially cleaved by plasmin or
CC protease specific antigen (PSA). The conjugate of the invention is useful
CC for treating a cell-surface protease-associated disease such as: cancer;
CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
CC wounds, circulatory disorders, dermatological disorders, rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
CC scarring from excimer laser surgery, scarring from glaucoma filtering
CC surgery, macular degeneration, cress syndromes, solid neoplasms, vascular
CC tumours, melanoma and Kaposi's sarcoma. The present DNA sequence encodes
CC a human cell surface protease.
XX
XX Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;

Query Match 95.2%; Score 1663.4; DB 10; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q 11 ATGAGAGGGGACACGCGGAAATGCATCTCCAGCAAGAACCTTCAGTGAAGCATCT 70
D 1 ATGAGAGGGGACACGCGGAAATGCATCTCCAGCAAGAACCTTCAGTGAAGCATCT 60
Q 71 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 130
D 61 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 120
Q 131 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 190
D 121 CCAAGCCAGGCATCTCCAGCTGGGACACCTCCAGCCCGGGCATCTCCAGCCAGGCATCT 180
Q 191 CCAAGCTGTACACTCCAGGCGGGGACATCTCCAGGCGGGGATCTCCAGCCAGGCATCT 250
D 181 CCAAGCTGTACACTCCAGGCGGGGACATCTCCAGGCGGGGATCTCCAGCCAGGCATCT 240
Q 251 CCAAGCCCGGCATCTCCGCTCTGGCATCACTTCCAGTCTTCATCCGCGAGTCACTCA 310
D 241 CCAAGCCCGGCATCTCCGCTCTGGCATCACTTCCAGTCTTCATCCGCGAGTCACTCA 300
Q 311 TCCGCGAGTCAAGCTCTGGTGAACAACCTCCCAACCAAGTGAACCTTTTGAACAACA 370
D 301 TCCGCGAGTCAAGCTCTGGTGAACAACCTCCCAACCAAGTGAACCTTTTGAACAACA 360
Q 371 CCAAGTGGGGGCTGTAACCATCCGATCATCTCTGCGCAGGTCAGACAGCAACGAGGCC 430
D 361 CCAAGTGGGGGCTGTAACCATCCGATCATCTCTGCGCAGGTCAGACAGCAACGAGGCC 420
Q 431 ACCAGGAGAGCCAGGTAAGAGCTCCCAAGTTCACTTGGCGGAGAGGCCAGAAAGCAG 490
D 421 ACCAGGAGAGCCAGGTAAGAGCTCCCAAGTTCACTTGGCGGAGAGGCCAGAAAGCAG 480
Q 491 CTACCGCTCATCGGGTCGTGCTCTCTCATTTGCTGCTGTTTGGCTCATATCTCTC 550
D 481 CTACCGCTCATCGGGTCGTGCTCTCTCATTTGCTGCTGTTTGGCTCATATCTCTC 540
Q 551 TTCCAGTTCTGGGAGGAGCAACAAGGATGATGAGAGAGAGAGAGAGCTGTCTCC 610
D 541 TTCCAGTTCTGGGAGGAGCAACAAGGATGATGAGAGAGAGAGAGAGAGCTGTCTCC 600
Q 611 AAGCAGCTGTCTGCTGTGAACGGGTGTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 670
D 601 AAGCAGCTGTCTGCTGTGAACGGGTGTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
Q 671 TGGGTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 730
D 661 TGGGTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720

Qy	731	TGGCTTCCCAATCTGTGAGAGCAACTGGAAATGACTCTCTCATAGAGAAAGCACTGCGCAGAG	790
Db	721	TGGCTTCCCAATCTGTGAGAGCAACTGGAAATGACTCTCTCATAGAGAAAGCACTGCGCAGAG	780
Qy	791	CTGGGTTTCGAGAGAGTGTCTCACCGGCAACCGAGGTTGGCCACAGGGATTTTGGCCAAAGC	850
Db	781	CTGGGTTTCGAGAGAGTGTCTCACCGGCAACCGAGGTTGGCCACAGGGATTTTGGCCAAAGC	840
Qy	851	TTCTCAATCTTTAGATATCAACTTCCACCATCCAGGAAAGCTTCCACAGGTTGATGCTT	910
Db	841	TTCTCAATCTTTAGATATCAACTTCCACCATCCAGGAAAGCTTCCACAGGTTGATGCTT	900
Qy	911	TCCCAAGCGGTATATCTCCCTCCAGTGTTCCTCCACTGGCCGATCTGAGGGCCATGACCGGGGGG	970
Db	901	TCCCAAGCGGTATATCTCCCTCCAGTGTTCCTCCACTGGCCGATCTGAGGGCCATGACCGGGGGG	960
Qy	971	ATCGTGGAGAGGAGCGCTGGCTGGCTCGGATGAGAAAGTGGCTTGGCAAGTAGTGTGCACTTC	1038
Db	961	ATCGTGGAGAGGAGCGCTGGCTGGCTCGGATGAGAAAGTGGCTTGGCAAGTAGTGTGCACTTC	1020
Qy	1031	GGACCAACCCACATCTGTGAGAGCAAGCTCATATGACGCCACAGTGGTCTCATGCGCC	1090
Db	1021	GGACCAACCCACATCTGTGAGAGCAAGCTCATATGACGCCACAGTGGTCTCATGCGCC	1080
Qy	1091	CACGTCTTCTTCGTGACCCCGGGAGAAAGTCTCTGAGAGGCTTGGAAAGTGTATCCCGGGCAC	1158
Db	1081	CACGTCTTCTTCGTGACCCCGGGAGAAAGTCTCTGAGAGGCTTGGAAAGTGTATCCCGGGCAC	1140
Qy	1151	AGCAACCTGCACAGTTGCTGAGGACGCTTCATTTGGCGAGATCATCATCAACAGCAAT	1210
Db	1141	AGCAACCTGCACAGTTGCTGAGGACGCTTCATTTGGCGAGATCATCATCAACAGCAAT	1200
Qy	1211	TACACCGATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGAC	1270
Db	1201	TACACCGATGAGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAAGCCCTGAC	1260
Qy	1271	CTGTCCGCTCAATTCACCCCTGCTTGCTTCCCTCCATGATGACAGACCTTTTAGCTCAAT	1330
Db	1261	CTGTCCGCTCAATTCACCCCTGCTTGCTTCCCTCCATGATGACAGACCTTTTAGCTCAAT	1320
Qy	1331	GAGACTGTCTGATATCACAGGCTTTTGGACAAGACCAAGGAGAGACGATGACAAGACATCCCC	1390
Db	1321	GAGACTGTCTGATATCACAGGCTTTTGGACAAGACCAAGGAGAGACGATGACAAGACATCCCC	1380
Qy	1391	TTCTCCGGAAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTCTGCTC	1458
Db	1381	TTCTCCGGAAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGACTTCTGCTC	1448
Qy	1451	TATGACAGTTATCTTACCCCAAGATGATATGTGTCTGTGGGACCTTCTGTGGGGCAGAGAC	1510
Db	1441	TATGACAGTTATCTTACCCCAAGATGATATGTGTCTGTGGGACCTTCTGTGGGGCAGAGAC	1500
Qy	1511	TCTGTGCCAGGAGAGACGCGGGGGGCTCTTGTCTGTGAGAGAAACACGCGTGGATACCTG	1570
Db	1501	TCTGTGCCAGGAGAGACGCGGGGGGCTCTTGTCTGTGAGAGAAACACGCGTGGATACCTG	1560
Qy	1571	GCAGGTGTCAACAGCTGGGAGCACAGGCTGTGGCCAGAGAAAACAACTGTGTGTATCAC	1630
Db	1561	GCAGGTGTCAACAGCTGGGAGCACAGGCTGTGGCCAGAGAAAACAACTGTGTGTATCAC	1620
Qy	1631	AAAGTGACAGAAATTTCTTCCCTGTGATTTTACACAGAAATGAGAGC	1675
Db	1621	AAAGTGACAGAAATTTCTTCCCTGTGATTTTACACAGAAATGAGAGC	1665

DE Human transmembrane serine protease (MTSP)-related cDNA #3.
XX
XX Human; transmembrane serine protease; MTSP; gene; 88;
KW cell surface protease; plasmin; prostate specific antigen; PSA;
KW proliferative disease; cell-surface protease-associated disease;
KW autoimmune disease; inflammatory disease; infectious disease;
KW endocrine disease; cancer; ocular disorder; cardiovascular disorder;
KW chronic inflammatory disease; wound; circulatory disorder;
KW dermatological disorder; restenosis; rheumatoid arthritis; psoriasis;
KW diabetic retinopathy; laser surgery; scarring;
KW glucose filtering surgery scarring; macular degeneration;
KW CRST syndrome; bacterial infection; viral disease; solid neoplasm;
KW vascular tumour; lung; colon; prostate; melanoma; Kaposi's sarcoma.
XX
OS Homo sapiens.
XX
XX US2004001801-A1.
XX
XX 01-JAN-2004.
XX
XX 23-MAY-2002; 2002US-00156214.
XX
XX 23-MAY-2002; 2002US-00156214.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Sempie JE, Vlaauk GP, Kemp SJ, Komandla M, Stew DV,
PI WPI: 2004-190126/18.
XX
XX P-PSDB; ADJ46919.
XX
XX Conjugate useful for treating e.g. cancer, cell-surface protease-
PT associated diseases, comprising a peptidic substrate or nucleic acid
PT substrate linked to a therapeutic agent through a linker.
XX
XX
XX
XX
XX Claim 9; SEQ ID NO 25; 361pp; English.
XX
XX The invention relates to a conjugate comprising a therapeutic agent and a
XX peptidic substrate or nucleic acid substrate linked to the agent
XX optionally through a linker or peptidic linker, where the peptidic
XX substrate is proteolytically cleaved by a cell surface protease or a
XX soluble, released or shed form conjugate to liberate the agent and the
XX conjugate is not substantially cleaved by plasmin or prostate specific
XX antigen (PSA). The conjugate is useful for treating a disease, which
XX involves administering a conjugate to a subject, where the disease is
XX preferably a proliferative disease or a cell-surface protease-associated
XX disease. The diseases include autoimmune diseases, inflammatory diseases,
XX infectious diseases and endocrine diseases. The conjugate is useful for
XX treating a cell-surface protease-associated disease, which involves
XX administering a conjugate comprising an agent and a peptidic substrate
XX a subject exhibiting symptoms of a cell-surface protease-associated
XX disorder, where the disease is selected from cancer, ocular disorders,
XX cardiovascular disorders, chronic inflammatory diseases, wounds,
XX circulatory disorders, dermatological disorders, restenosis, rheumatoid
XX arthritis, psoriasis, diabetic retinopathies, scarring from laser
XX surgery, scarring from glaucoma filtering surgery, macular degeneration,
XX CRST syndrome, bacterial infections, viral diseases, solid neoplasms and
XX vascular tumours such as lung cancer, colon cancer, prostate cancer,
XX melanoma and Kaposi's sarcoma. This sequence represents cDNA encoding a
XX transmembrane serine protease (MTSP)-related polypeptide of the
XX invention.
XX
XX Sequence 2066 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;
XX

Db 61 CCAAGCCAGGATCTCCAGCTGGGACA CTTCAAGCCGGGATCTCCAGCCAGGACATCT 120
Qy 131 CCAAGCCAGGATCTCCAGCTGGGACA CTTCCGGGCGGGGATCTCCAGCCAGGACATCT 190
Db 121 CCAAGCCAGGATCTCCAGCTGGGACA CTTCCGGGCGGGGATCTCCAGCCAGGACATCT 180
Qy 191 CCAAGCTGTACATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGACATCT 250
Db 181 CCAAGCTGTACATCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGACATCT 240
Qy 251 CCAAGCCGGGATCTCCGGGCTCTGGGATCACTTTCCAGGTCCTGATCCGGAGGATCACTCA 310
Db 241 CCAAGCCGGGATCTCCGGGCTCTGGGATCACTTTCCAGGTCCTGATCCGGAGGATCACTCA 300
Qy 311 TCCGCGAGGTCAGCTCGGTGAACA CTTCCCAACAGAGTACCTTTGTAAGACA 370
Db 301 TCCGCGAGGTCAGCTCGGTGAACA CTTCCCAACAGAGTACCTTTGTAAGACA 360
Qy 371 CCAAGTGGGGGCTGATCCATCCGATCTCTGCTCCAGGTAGAGACAGCAAGCAAGGAGCC 430
Db 361 CCAAGTGGGGGCTGATCCATCCGATCTCTGCTCCAGGTAGAGACAGCAAGCAAGGAGCC 420
Qy 431 ACCAGGAGAGCCAGGATACAGGCTGCGCAAGTTCACTGGCGGAGAGCCAGAGAGAG 490
Db 421 ACCAGGAGAGCCAGGATACAGGCTGCGCAAGTTCACTGGCGGAGAGCCAGAGAGAG 480
Qy 491 CTACCGCTCATCGGGTGGGTCTCTCTCTCATGTCCTGCTGCTGCTCATCTCTC 550
Db 481 CTACCGCTCATCGGGTGGGTCTCTCTCTCATGTCCTGCTGCTGCTCATCTCTC 540
Qy 551 TTCCAGTTTGGGAGGGGCAACAAGGATCAGGATCAAGAGGAGAGGAGAGGATGTCCTC 610
Db 541 TTCCAGTTTGGGAGGGGCAACAAGGATCAGGATCAAGAGGAGAGGAGAGGATGTCCTC 600
Qy 611 AAGCAGCTGTTCTGCTGTGACGAGGAGTGGAGCTGCAAGCTGAGAGAGTGAACGACTGGGC 670
Db 601 AAGCAGCTGTTCTGCTGTGACGAGGAGTGGAGCTGCAAGCTGAGAGAGTGAACGACTGGGC 660
Qy 671 TGCCTGAGGTTTGACTGGGACAAGTCTCTGCTTAAATCTA CTTGGGTCCTCCCATCAG 730
Db 661 TGCCTGAGGTTTGACTGGGACAAGTCTCTGCTTAAATCTA CTTGGGTCCTCCCATCAG 720
Qy 731 TGGCTTCCATCTGTAGAGAGAACTGGAATCTCTACTAGAGAAAGAACTGTCAGAG 790
Db 721 TGGCTTCCATCTGTAGAGAGAACTGGAATCTCTACTAGAGAAAGAACTGTCAGAG 780
Qy 791 CTGGGTTTCGAGAGTGTCACTCGGACAACGAGGTTGCCACAGGAGATTTTGCCAAAGC 850
Db 781 CTGGGTTTCGAGAGTGTCACTCGGACAACGAGGTTGCCACAGGAGATTTTGCCAAAGC 840
Qy 851 TTCTCAATCTTGAATACAACTCCACATCCAGGAAAGCTTCA CAGGTCGAAATGCCCT 910
Db 841 TTCTCAATCTTGAATACAACTCCACATCCAGGAAAGCTTCA CAGGTCGAAATGCCCT 900
Qy 911 TCCGAGGAGTATATCTCCCTCAAGTTCCTCACTGCCGAGCTGAGAGGCACTAACCGGCGG 970
Db 901 TCCGAGGAGTATATCTCCCTCAAGTTCCTCACTGCCGAGCTGAGAGGCACTAACCGGCGG 960
Qy 971 ATGCTGGAGGAGGAGGCTCGGCTCGGATAGCAAGTGGCTTGG CAAAGTGAAGTCACTTC 1030
Db 961 ATGCTGGAGGAGGAGGCTCGGCTCGGATAGCAAGTGGCTTGG CAAAGTGAAGTCACTTC 1020
Qy 1031 GGCACCAACCAATCTGTGAGGACGCTCATYTAGACGCCAG GAGGTGCTCATGCGGCC 1090
Db 1021 GGCACCAACCAATCTGTGAGGACGCTCATYTAGACGCCAG GAGGTGCTCATGCGGCC 1080
Qy 1091 CACTGCTTCTCGTGAACCGGAGAAAGGCTCTGAGAGGCTGGA GAGGTGTAAGCGGAGACC 1150
Db 1081 CACTGCTTCTCGTGAACCGGAGAAAGGCTCTGAGAGGCTGGA GAGGTGTAAGCGGAGACC 1140
Qy 1151 AGCAACCTGACCAAGTTGCTGAGAGAGCTCATYTGCGAGAT CATCAAGCAAT 1210

Db 1141 AGCAACCTGACCAAGTTGCTGAGAGAGCTTCAATGCGGATGATCATCAAGCAAT 1200
Qy 1211 TACACCGATAGAGAGAGCACTATGATGAGCCCTCAATGAGGCTGATCAAGCCCTGACC 1270
Db 1201 TACACCGATAGAGAGAGCACTATGATGAGCCCTCAATGAGGCTGATCAAGCCCTGACC 1260
Qy 1271 CTGTCCGCTCAATCCACCTGCTTGGCTTCCCATGATGAGAGACCTTTAGCTCAAT 1330
Db 1261 CTGTCCGCTCAATCCACCTGCTTGGCTTCCCATGATGAGAGACCTTTAGCTCAAT 1320
Qy 1331 GAGACTGCTGATACAGGCTTTGGCAAGACAGGAGACAGATGACAGATCAATCCCC 1390
Db 1321 GAGACTGCTGATACAGGCTTTGGCAAGACAGGAGACAGATGACAGATCAATCCCC 1380
Qy 1391 TTCTCCGAGGAGTGCAGGTCATCTGATGACTTCAAGAAATGCAATGACTTGGTC 1450
Db 1381 TTCTCCGAGGAGTGCAGGTCATCTGATGACTTCAAGAAATGCAATGACTTGGTC 1440
Qy 1451 TATGACAGTTACCTTACCCCAAGGATGATGTCGAGGACCTT CAGGAGGAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGGATGATGTCGAGGACCTT CAGGAGGAGAGAC 1500
Qy 1511 TCTGCGAGGAGAGACAGCGGGGCTCTTGTCTGTGAGAGAA CAACTGCTGATCTG 1570
Db 1501 TCTGCGAGGAGAGACAGCGGGGCTCTTGTCTGTGAGAGAA CAACTGCTGATCTG 1560
Qy 1571 GCAGGTTCACCAAGTGGGAGACAGGCTGTGGCCAGAGAA CAACTGCTGATCAACC 1630
Db 1561 GCAGGTTCACCAAGTGGGAGACAGGCTGTGGCCAGAGAA CAACTGCTGATCAACC 1620
Qy 1631 AAAGTGACAGAAAGTTCTTCCCTGATTTTACAGCAAGTGA GAGAGC 1675
Db 1621 AAAGTGACAGAAAGTTCTTCCCTGATTTTACAGCAAGTGA GAGAGC 1665

RESULT 13
ADY50053
ID ADY50053 standard; DNA; 2067 BP.
XX
AC ADY50053;
XX
DT 19-MAY-2005 (first entry)
XX
DE Human endothelias-2L DNA.
XX
XX Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;
KW Ophthalmological; Cardiovascular-Gen.; Vulnery; Antiinflammatory;
KW Vasoconstrictive; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
KW rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
KW dermatological disease; cancer; neoplasm; endothelias-2; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2005019270-A2.
XX
PD 03-MAR-2005.
XX
PF 12-AUG-2004; 2004WO-US026148.
XX
PR 14-AUG-2003; 2003US-0495005P.
XX
PR 14-NOV-2003; 2003US-0520164P.
XX
PA (DYAX-) DYAX CORP.
XX
PA (DEN-) DENREON CORP.
XX
PI Madison EL, Nixon A;
XX
XX WPI: 2005-202609/21.
XX
XX P-PSDB; ADY50054.
XX
PT Novel protein capable of inhibiting endothelias-2, useful for treating
or preventing angiogenesis related disorder e.g. cancer.
XX

PS Disclosure; SEQ ID NO 1; 157bp; English.
XX The invention relates to an isolated protein (I) which comprises a heavy
CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
CC immunoglobulin variable domain sequence, where the first and second
CC immunoglobulin variable domain sequences from an antigen binding site
CC that specifically binds to human endothelialase-2. (I) is useful for
CC detecting an endothelialase or endothelialase activity in a sample, for
CC modulating an activity of an ET2-expressing cell, for modulating
CC proteolysis, for killing or inhibiting growth of a cell, for detecting
CC endothelialase in a subject, for modulating endothelialase activity in a
CC subject, for treating or preventing a disorder characterized by unwanted
CC angiogenesis in a subject. The disorder is chosen from rheumatoid
CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
CC pterygii recurrence, scarring excimer laser surgery and glaucoma
CC filtering surgery, cardiovascular disorders, chronic inflammatory
CC disorders, wound repair, circulatory disorders, crest syndromes,
CC dermatological disorders and cancers. The present sequence represents the
CC human endothelialase-2L DNA.
XX
SQ Sequence 2067 BP; 463 A; 604 C; 540 G; 460 T; 0 U; 0 Other;
Query Match 95.2%; Score 1663.4; DB 14; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 11 ATGAGAGAGGACACCGAGATGATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
DB 1 ATGAGAGAGGACACCGAGATGATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 60
QY 71 CCAAGCCCAAGCATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 130
DB 61 CCAAGCCCAAGCATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 120
QY 131 CCAAGCCCAAGCATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 190
DB 121 CCAAGCCCAAGCATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGGATCT 180
QY 191 CCAAGCTGTACCTCCAGAGCCGGGATCTTCAGAGCCGGGATCTCCAGCCAGGATCT 250
DB 181 CCAAGCTGTACCTCCAGAGCCGGGATCTTCAGAGCCGGGATCTCCAGCCAGGATCT 240
QY 251 CCAAGCCCGGATCTCCAGCTGGGATCTTCAGAGCCGGGATCTTCAGAGCCAGGATCTCA 310
DB 241 CCAAGCCCGGATCTCCAGCTGGGATCTTCAGAGCCGGGATCTTCAGAGCCAGGATCTCA 300
QY 311 TCCGCGAGGTAGAGCTGGGATCAACCTCCCAACAGAGTACTCTTGTATGACAA 370
DB 301 TCCGCGAGGTAGAGCTGGGATCAACCTCCCAACAGAGTACTCTTGTATGACAA 360
QY 371 CCAAGTGGGGGCTGATCCATCCATCTCTCCAGAGTACAGCAACAGAGGCC 430
DB 361 CCAAGTGGGGGCTGATCCATCCATCTCTCCAGAGTACAGCAACAGAGGCC 420
QY 431 ACCAGGAGAGCCAGAGTACAGAGCTCCCAAGTTCACTGGCGGAGAGGCCAGAGAG 490
DB 421 ACCAGGAGAGCCAGAGTACAGAGCTCCCAAGTTCACTGGCGGAGAGGCCAGAGAG 480
QY 491 CTACCGCTCATCGGGTGGTCTCTCTCTCATGCGCCCTGGTGGTTCATATCTCTC 550
DB 481 CTACCGCTCATCGGGTGGTCTCTCTCTCATGCGCCCTGGTGGTTCATATCTCTC 540
QY 551 TTCAGATTCTGGAGAGGACACAGAGATCAGGTACAAAGAGAGAGAGAGTGTCTCC 610
DB 541 TTCAGATTCTGGAGAGGACACAGAGATCAGGTACAAAGAGAGAGAGAGTGTCTCC 600
QY 611 AAGCAGCGTGTTCGCTGTACAGGGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 670
DB 601 AAGCAGCGTGTTCGCTGTACAGGGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
QY 671 TGGGTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 730
DB 661 TGGGTGAGGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720

QY 731 TGGGTTCCTCATCTGTACAGCAACTGGATGACTCTTACTCAGAGAGAGTGTCCAGAG 790
DB 721 TGGGTTCCTCATCTGTACAGCAACTGGATGACTCTTACTCAGAGAGAGTGTCCAGAG 780
QY 791 CTGGGTTCAGAGAGTGTCTACCCGAGCAACGAGGTTGCCCAAGAGATTTTGGCAACAG 850
DB 781 CTGGGTTCAGAGAGTGTCTACCCGAGCAACGAGGTTGCCCAAGAGATTTTGGCAACAG 840
QY 851 TTCCTCATCTTGAATGATCACTCCATCATGAGGAAAGCTCCACAGGTGAAATGCCCT 910
DB 841 TTCCTCATCTTGAATGATCACTCCATCATGAGGAAAGCTCCACAGGTGAAATGCCCT 900
QY 911 TCCAGGAGTATATCTCTCCAGTGTTCCTCATCTGCGAGTGAAGGACATGACCGGAGCG 970
DB 901 TCCAGGAGTATATCTCTCCAGTGTTCCTCATCTGCGAGTGAAGGACATGACCGGAGCG 960
QY 971 ATGCTGGAGAGGAGGCTGAGCTTGGATAGCAAGTGGCTTGGCAAGTGAATCTGCACTTC 1030
DB 961 ATGCTGGAGAGGAGGCTGAGCTTGGATAGCAAGTGGCTTGGCAAGTGAATCTGCACTTC 1020
QY 1031 GGCACCAACCCATCTGTGAGAGGAGCTGATGAGCCGAGTGGGAGTCACTGAGCCGC 1090
DB 1021 GGCACCAACCCATCTGTGAGAGGAGCTGATGAGCCGAGTGGGAGTCACTGAGCCGC 1080
QY 1091 CACTGCTTCTTGTGACCCGAGAGAGGCTCTGAGAGGCTGGAAGTGTACCGCGGAGCC 1150
DB 1081 CACTGCTTCTTGTGACCCGAGAGAGGCTCTGAGAGGCTGGAAGTGTACCGCGGAGCC 1140
QY 1151 AGCAACTGACCAAGTGTGCTGAGAGGAGCTTCCATTCGAGATCATCATCAAGCAAT 1210
DB 1141 AGCAACTGACCAAGTGTGCTGAGAGGAGCTTCCATTCGAGATCATCATCAAGCAAT 1200
QY 1211 TACACGATGAGAGAGAGCTATGATGATGAGCTTCAAGGAGCTGTCCAAGCCCTGACC 1270
DB 1201 TACACGATGAGAGAGAGCTATGATGATGAGCTTCAAGGAGCTGTCCAAGCCCTGACC 1260
QY 1271 CTGTCCGCTCAATCCACCTGCTTGGCTTCCATGATGAGAGCACTTTAGCTCAAT 1330
DB 1261 CTGTCCGCTCAATCCACCTGCTTGGCTTCCATGATGAGAGCACTTTAGCTCAAT 1320
QY 1331 GAGACTGCTGATCAAGGCTTGGCAAGACAGGAGAGACATGACAAAGATCTCCGCC 1390
DB 1321 GAGACTGCTGATCAAGGCTTGGCAAGACAGGAGAGACATGACAAAGATCTCCGCC 1380
QY 1391 TTCTCCGAGAGGAGGAGTCAATCTCATGAGCTTCAAGAAATGCAATGACTACTTGTGTC 1450
DB 1381 TTCTCCGAGAGGAGGAGTCAATCTCATGAGCTTCAAGAAATGCAATGACTACTTGTGTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTGGGGGAGAGAC 1510
DB 1441 TATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGTGGGGGAGAGAC 1500
QY 1511 TCTGCGCAGGAGAGACAGCGGGGCTCTTGTCTGTATGAGCAAGCAACCGCTGTACTCTG 1570
DB 1501 TCTGCGCAGGAGAGACAGCGGGGCTCTTGTCTGTATGAGCAAGCAACCGCTGTACTCTG 1560
QY 1571 GCAAGTGTCAACAGCTGGGAGCAAGGCTGTGGCCAGAGAAACAACTGTGTGTACAC 1630
DB 1561 GCAAGTGTCAACAGCTGGGAGCAAGGCTGTGGCCAGAGAAACAACTGTGTGTACAC 1620
QY 1631 AAAGTGAAGAGTCTTCTCTGATTTTACAGCAAGTGAAGAC 1675
DB 1621 AAAGTGAAGAGTCTTCTCTGATTTTACAGCAAGTGAAGAC 1665
RESULT 14
ADH17413
ID ADH17413 standard; cDNA; 2432 BP.
XX
XX ADH17413;
AC
XX 11-MAR-2004 (first entry)

Human NOV12a cDNA - SEQ ID 103.

XX Human NOV12a cDNA - SEQ ID 103.
DE
XX
XX NOVX: antidiabetic; anorectic; cardiant; hypotensive;
KM antiatherosclerotic; anorectic; vincicide; antibacterial;
KM protoascide; nootropic; neuroprotective; antiparkinsonian;
KM anticonvulsant; osteopathic; antiarthritis; antiinflammatory;
KM dematological; antiasthmatic; antidiabetic; metabolic; diabetes;
KM obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KM atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's
KM epilepsy; immune; osteoarthritis; haemopoietic;
KM inflammatory skin disorder; asthma; dyslipidaemia;
KM cell differentiation; proliferation; haemopoiesis; neurogenesis;
KM angiogenesis; gene therapy; chromosome mapping; wound healing;
KM pharmacogenomic; human; ss; gene.

OS Homo sapiens.

PN WO2003093432-A2.

PD 13-NOV-2003 .

PF 02-MAY-2003; 2003WO-US013690.

PR	02-MAY-2002;	2002US-037733.1P
PR	08-MAY-2002;	2002US-037673.0P
PR	24-MAY-2002;	2002US-038307.5P
PR	29-MAY-2002;	2002US-038404.4P
PR	30-MAY-2002;	2002US-038421.5P
PR	30-MAY-2002;	2002US-038429.6P
PR	30-MAY-2002;	2002US-038432.9P
PR	30-MAY-2002;	2002US-038433.2P
PR	30-MAY-2002;	2002US-038435.2P
PR	31-MAY-2002;	2002US-038521.1P
PR	02-JUL-2002;	2002US-039333.3P
PR	09-AUG-2002;	2002US-040215.4P
PR	09-AUG-2002;	2002US-040217.1P
PR	09-AUG-2002;	2002US-040220.4P
PR	09-AUG-2002;	2002US-040220.5P
PR	22-AUG-2002;	2002US-040517.5P
PR	27-AUG-2002;	2002US-040613.9P
PR	23-SEP-2002;	2002US-041295.4P
PR	30-SEP-2002;	2002US-041497.5P
PR	07-OCT-2002;	2002US-041661.1P
PR	24-OCT-2002;	2002US-042085.1P
PR	31-OCT-2002;	2002US-042254.7P
PR	01-MAY-2003;	2003US-004283.75

(CURA-) CURAGEN CORP.

PI Alvarez, Anderson DW, Boldo FL, Catterton E, Edinger SR;
PI Fernandez ER, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
PI Kendra R, Li L, Macdougall JR, Padigar M, Paturdian M,
PI Peterson JD, Rastelli L, Shmets RA, Sylek KA, Stone DJ,
PI Vernet CM, Voss E, Zhong M;

DR WPI; 2004-053040/05.

DR P-PSDB; ADH17414.

PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections

PS Claim 20; SEQ ID NO 103; 478bp; English.

The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates antidiabetic, anorectic, cardiatic, hypotensive, antihypertensive, anorectic, virocidic, antibacterial, fungicide, protozoocidal, nootropic, neuroprotective, antiparkinsonian, anticonvulsant, osteoplastic, antiallergic, antiinflammatory, dematocidal, antiepileptic and antipapillary activities. The polypeptide, nucleic acid molecules and antibodies be useful in the manufacture of a medicament for treating metabolic

CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic.
CC antiinflammatory, dermatological, antistaphylococcal and antilipemic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOXV substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOXV cDNA of the invention.

XX
SQ Sequence 2432 BP; 568 A; 701 C; 637 G; 526 T; 0 U; 0 Other;

Query Match 94.0%; Score 1643.8; DB 12; Length 2432;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 CTCAGAGCCATGAGAGGAGGACACCGGAGATGATCTCCAGCAAGAACACCTTCAGC 60
Db CTGAGAGACCATGAGAGGAGGACACCGGAGATGATCTCCAGCAAGAACACCTTCAGC 161
QY 61 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 120
Db TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 221
QY 121 CCAGGCAATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 180
Db CCAGGCAATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 281
QY 222 CCAGGCAATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 281
Db 181 CCAGGCAATCTCCAGCTGGTACACCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 228
QY 282 CCAGGCAATCTCCAGCTGGTACACCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 341
Db 282 CCAGGCAATCTCCAGCTGGTACACCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 341
QY 229 ---GGCATCTCCAGCCAGGATCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 285
Db 342 CCAGGCAATCTCCAGCCAGGATCTCCAGCCGGGATCTCCAGCCGGGATCTCCAGC 401
QY 286 CAGGTCTCATCCGGAGGATCATCATCCAGGATCTCCAGCCGGGATCTCCAGC 345
Db 402 CAGGTCTCATCCGGAGGATCATCATCCAGGATCTCCAGCCGGGATCTCCAGC 461
QY 346 CAGAGTGTACTTGTAGAGCAACCAAGTGGGGGCTGTACCATCCGATCATCTCTGC 405
Db 462 CAGAGTGTACTTGTAGAGCAACCAAGTGGGGGCTGTACCATCCGATCATCTCTGC 521
QY 406 CAGGTCAAGACCAAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 465
Db 522 CAGGTCAAGACCAAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581
QY 466 CAGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 525
Db 582 CAGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 641
QY 526 CAGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 585
Db 642 CAGGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 701
QY 586 CAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 645
Db 702 CAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 761
QY 646 CAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 705

Db 762 CAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 821
QY 706 AATCTACTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 765
Db 822 AATCTACTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 881
QY 766 CTACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 825
Db 882 CTACTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 941
QY 826 TGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 885
Db 942 TGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1001
QY 886 AAGCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 945
Db 1002 AAGCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1061
QY 946 CGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1005
Db 1062 CGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1121
QY 1006 GCTTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065
Db 1122 GCTTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1181
QY 1066 CGCCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1125
Db 1182 CGCCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1241
QY 1126 GGGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1185
Db 1242 GGGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1301
QY 1186 TGCCAGATCATCATCAACAGCAATTAACAGATGAGAGGAGGAGGAGGAGGAGGAGG 1245
Db 1302 TGCCAGATCATCATCAACAGCAATTAACAGATGAGAGGAGGAGGAGGAGGAGGAGG 1361
QY 1246 CATCGGCTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1305
Db 1362 CATCGGCTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1421
QY 1306 GCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1365
Db 1422 GCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1481
QY 1366 GAGAGCAGATGAGCAAGACATCCCTCTCTCCGAGAGGAGGAGGAGGAGGAGGAGG 1425
Db 1482 GAGAGCAGATGAGCAAGACATCCCTCTCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGG 1541
QY 1426 CAGGAATGCAATGATCTTGTGTATGACAGTTACCTTAACCAAGATGATGTGTGC 1485
Db 1542 CAGGAATGCAATGATCTTGTGTATGACAGTTACCTTAACCAAGATGATGTGTGC 1601
QY 1486 TGGGGAGCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1545
Db 1602 TGGGGAGCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1661
QY 1546 TGAAGCAGAGCAACCGCTGTATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1605
Db 1662 TGAAGCAGAGCAACCGCTGTATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1721
QY 1606 GAGAAACAAACCTGGTGTGTACACCAAGTGAAGAAAGTTCTTCCGTGATTTACAGCAA 1665
Db 1722 GAGAAACAAACCTGGTGTGTGTACACCAAGTGAAGAAAGTTCTTCCGTGATTTACAGCAA 1781
QY 1666 GATGAGG 1672
Db 1782 GATGAGG 1788

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 17, 2006, 13:53:14 / Search time 567 Seconds
(without alignments)
5768.427 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 1748
Sequence: 1 ctccagagcattctcagagaggg.....ggtctgctgacccagagaa 1748

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /EMC_Celerra_SIDS3/prodata/2/ina/1.COMB.seq: *
2: /EMC_Celerra_SIDS3/prodata/2/ina/5.COMB.seq: *
3: /EMC_Celerra_SIDS3/prodata/2/ina/6A.COMB.seq: *
4: /EMC_Celerra_SIDS3/prodata/2/ina/6B.COMB.seq: *
5: /EMC_Celerra_SIDS3/prodata/2/ina/7.COMB.seq: *
6: /EMC_Celerra_SIDS3/prodata/2/ina/H.COMB.seq: *
7: /EMC_Celerra_SIDS3/prodata/2/ina/RE.COMB.seq: *
8: /EMC_Celerra_SIDS3/prodata/2/ina/PP.COMB.seq: *
9: /EMC_Celerra_SIDS3/prodata/2/ina/BACKFILE1.seq: *
10: /EMC_Celerra_SIDS3/prodata/2/ina/BACKFILE1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1748	100.0	1748	US-09-879-792-11	Sequence 11, Appl
2	1258.6	72.0	1434	US-10-177-661-1	Sequence 1, Appl
3	1062.6	60.8	1341	US-10-177-661-3	Sequence 3, Appl
4	1053	60.0	1230	US-09-879-792-35	Sequence 35, Appl
5	709.4	40.6	834	US-09-879-792-38	Sequence 28, Appl
6	440.8	25.2	678	US-09-879-792-30	Sequence 30, Appl
7	411.6	23.5	614	US-09-879-792-33	Sequence 33, Appl
8	348	19.9	402	US-09-879-792-1	Sequence 1, Appl
9	335.8	19.2	621	US-09-879-792-29	Sequence 29, Appl
10	333.2	19.1	610	US-09-280-116-22	Sequence 22, Appl
11	285	16.3	285	US-09-879-792-2	Sequence 2, Appl
12	276.4	15.8	286	US-09-879-792-5	Sequence 5, Appl
13	265.4	15.2	751	US-09-879-792-34	Sequence 34, Appl
14	193.2	11.1	2038	US-09-008-271A-18	Sequence 18, Appl
15	193.2	11.1	2038	US-09-068-415-18	Sequence 18, Appl
16	193.2	11.1	2063	US-10-012-231A-274	Sequence 274, Appl
17	193.2	11.1	2063	US-10-015-389A-274	Sequence 274, Appl
18	193.2	11.1	2063	US-10-006-768A-274	Sequence 274, Appl
19	193.2	11.1	2063	US-10-015-671A-274	Sequence 274, Appl
20	193.2	11.1	2063	US-10-015-393A-274	Sequence 274, Appl
21	193.2	11.1	2063	US-10-011-833A-274	Sequence 274, Appl
22	193.2	11.1	2063	US-10-006-041A-274	Sequence 274, Appl
23	193.2	11.1	2063	US-10-012-064A-274	Sequence 274, Appl

24	193.2	11.1	2063	US-10-015-392A-274	Sequence 274, Appl
25	193.2	11.1	2063	US-10-011-795B-274	Sequence 274, Appl
26	193.2	11.1	2063	US-10-015-386A-274	Sequence 274, Appl
27	193.2	11.1	2063	US-10-012-121A-274	Sequence 274, Appl
28	193.2	11.1	2063	US-10-006-485A-274	Sequence 274, Appl
29	193.2	11.1	2063	US-10-006-746A-274	Sequence 274, Appl
30	193.2	11.1	2063	US-10-012-752A-274	Sequence 274, Appl
31	193.2	11.1	2063	US-10-017-253A-274	Sequence 274, Appl
32	193.2	11.1	2063	US-10-015-519A-274	Sequence 274, Appl
33	193.2	11.1	2063	US-10-015-715A-274	Sequence 274, Appl
34	193.2	11.1	2063	US-10-007-236A-274	Sequence 274, Appl
35	193.2	11.1	2079	US-09-656-002-1	Sequence 1, Appl
36	193.2	11.1	2079	US-09-851-588-5	Sequence 1, Appl
37	193.2	11.1	2081	US-09-851-588-7	Sequence 7, Appl
38	191.6	11.0	2121	US-09-607-745-1	Sequence 1, Appl
39	190.8	10.9	2440	US-09-949-016-5210	Sequence 5210, Appl
40	190.8	10.9	2440	US-09-949-016-5211	Sequence 5211, Appl
41	190.8	10.9	2440	US-09-949-016-5212	Sequence 5212, Appl
42	190.4	10.9	1305	US-10-030-688-1	Sequence 1, Appl
43	189.6	10.8	1479	US-10-030-688-3	Sequence 3, Appl
44	185.2	10.6	688	US-09-879-792-32	Sequence 32, Appl
45	182.8	10.5	2412	US-09-999-833A-68	Sequence 68, Appl

ALIGNMENTS

```
RESULT 1
US-09-879-792-11
; Sequence 11, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Genrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1748
; TYPE: DNA
; ORGANISM: Homo sapiens
Query Match      100.0%; Score 1748; DB 3; Length 1748;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CTCAGAGACATCTGAGAGGAGGACACGCGGAAATGATCTTCACAGACAGAACCTTCAGC 60
1 CTCAGAGACATCTGAGAGGAGGACACGCGGAAATGATCTTCACAGACAGAACCTTCAGC 60
1 CTGAGAGACATCTGAGAGGAGGACACGCGGAAATGATCTTCACAGACAGAACCTTCAGC 60
61 TGGAGCATCTTCACAGCCAGGATCTTCAGCTGGAGACCTTCAGCGGGGATCTTCAGC 120
61 TGGAGCATCTTCACAGCCAGGATCTTCAGCTGGAGACCTTCAGCGGGGATCTTCAGC 120
61 TGGAGCATCTTCACAGCCAGGATCTTCAGCTGGAGACCTTCAGCGGGGATCTTCAGC 120
121 CCAGGACATCTTCAGCCAGGATCTTCAGCTGGAGACCTTCAGCGGGGATCTTCAGC 180
121 CCAGGACATCTTCAGCCAGGATCTTCAGCTGGAGACCTTCAGCGGGGATCTTCAGC 180
181 CCAGGACATCTTCAGCTGGATCTTCAGGCGGAGATCTTCAGCGGGGATCTTCAGC 240
```

Db 181 CCAGGCACTTCAGCTGTGATACCTCCAGCCGCGGCACTCCAGGCGCGGCACTCCAGC 240
Qy 241 CCAGGCACTTCAGGAGCCGCGGCACTCCGCGGCTCTGGCACTCACTTTCAGAGTCTCTACCG 300
Db 241 CCAGGCACTTCAGGAGCCGCGGCACTTCGCGGCTCTGGCACTCACTTTCAGAGTCTCTACCG 300
Qy 301 CAGGTCATCATCCGCGAGGTCAGGCTCGGTGACAACCTTCCCAACGAGGTGTAACCTTGT 360
Db 301 CAGGTCATCATCCGCGAGGTCAGGCTCGGTGACAACCTTCCCAACGAGGTGTAACCTTGT 360
Qy 361 TAGAGCAACACCACTGGGGGCTGTATCCCATCCGATCATCTCTCCGAGGTGAGACCAAC 420
Db 361 TAGAGCAACACCACTGGGGGCTGTATCCCATCCGATCATCTCTCCGAGGTGAGACCAAC 420
Qy 421 AACCAAGGAGCCACAGGAGAGCCAGGATCAGAGCTGCGCAAGTTCACTTGGCGGAGG 480
Db 421 AACCAAGGAGCCACAGGAGAGCCAGGATCAGAGCTGCGCAAGTTCACTTGGCGGAGG 480
Qy 481 CCAGAGAGGCTACCGCTCATCGGGTGCCTCTCTCTCATTTGCTGTGTTGCT 540
Db 481 CCAGAGAGGCTACCGCTCATCGGGTGCCTCTCTCTCTCATTTGCTGTGTTGCT 540
Qy 541 CATATCTCTCTCAGATTCCTGAGGAGGAGGAGGAGGATCAGAGTCAAGAGAGAGGAG 600
Db 541 CATATCTCTCTCAGATTCCTGAGGAGGAGGAGGAGGATCAGAGTCAAGAGAGAGGAG 600
Qy 601 GAGCTGTCCAGAGACGCTGTTCCTGTGACGCGGTTGTGATCTGCAAGCTGGAAGAGTA 660
Db 601 GAGCTGTCCAGAGACGCTGTTCCTGTGACGCGGTTGTGATCTGCAAGCTGGAAGAGTA 660
Qy 661 CGAGCTGGGCTGCGTGAAGTTGACTGAGGACAAGTCTGTGTTAAATCTACTCTGAGTC 720
Db 661 CGAGCTGGGCTGCGTGAAGTTGACTGAGGACAAGTCTGTGTTAAATCTACTCTGAGTC 720
Qy 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATACCTCTCTCTCTCTCTCT 780
Db 721 CTCCCATCAGTGGCTTCCCATCTGTAGAGAGCACTGGAATACCTCTCTCTCTCTCTCT 780
Qy 781 CTGCGCAGAGCTGGGTTTGAAGAGTGTCAACGAGGAGGAGGTTGCCACAGGATTT 840
Db 781 CTGCGCAGAGCTGGGTTTGAAGAGTGTCAACGAGGAGGAGGTTGCCACAGGATTT 840
Qy 841 TGCCAACAGCTTCTCAATCTTGAATACACTCCACATCCAGAGAAAGCTTCCACAGGTC 900
Db 841 TGCCAACAGCTTCTCAATCTTGAATACACTCCACATCCAGAGAAAGCTTCCACAGGTC 900
Qy 901 TGAATGCCCTTCCAGCGGTATATCTCTCCAGTGTCCCACTGCGGACTGAGGGCCAT 960
Db 901 TGAATGCCCTTCCAGCGGTATATCTCTCCAGTGTCCCACTGCGGACTGAGGGCCAT 960
Qy 961 GACCGGCGGATCGTGGAGAGGCGCTGCGCTCGGATGAGCAAGTGGCTTGGCAAGTAG 1020
Db 961 GACCGGCGGATCGTGGAGAGGCGCTGCGCTCGGATGAGCAAGTGGCTTGGCAAGTAG 1020
Qy 1021 TCTGCACTTGGGACCAACCACTCTGTGAGGACAAGCTCATTTGACGCGCCAGTGGTCT 1080
Db 1021 TCTGCACTTGGGACCAACCACTCTGTGAGGACAAGCTCATTTGACGCGCCAGTGGTCT 1080
Qy 1081 CACTGCGCGCCCACTGCTTCTTCCGTAACCGGAGAAAGTCTTGAAGGCTGGAAGTGT 1140
Db 1081 CACTGCGCGCCCACTGCTTCTTCCGTAACCGGAGAAAGTCTTGAAGGCTGGAAGTGT 1140
Qy 1141 CGCGGAGCAGCAACCTGTCACAGTTGCTGAGGAGCTCTCATTTGCGGAGATCATAT 1200
Db 1141 CGCGGAGCAGCAACCTGTCACAGTTGCTGAGGAGCTCTCATTTGCGGAGATCATAT 1200
Qy 1201 CAACAGCAATTACACCGATGAGAGAGAGACTATGACATGCGCTCATGCGGCTGTCCAA 1260
Db 1201 CAACAGCAATTACACCGATGAGAGAGAGACTATGACATGCGCTCATGCGGCTGTCCAA 1260
Qy 1261 GCGCCTGACCTGTTCGCTGATCATCCAGCTGCTTCCCTCCCATGATGAGACAGCTT 1320
Db 1261 GCGCCTGACCTGTTCGCTGATCATCCAGCTGCTTCCCTCCCATGATGAGACAGCTT 1320

Qy 1321 TAGCCTCAATGAGACCTGTGTGATCAGAGGCTTTGGCAAGACAGAGGAGACAGATGACAA 1380
Db 1321 TAGCCTCAATGAGACCTGTGTGATCAGAGGCTTTGGCAAGACAGAGGAGACAGATGACAA 1380
Qy 1381 GACATCCCGCTTCTCCGAGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA 1440
Db 1381 GACATCCCGCTTCTCCGAGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGCAATGA 1440
Qy 1441 CTACTGTGTATGAGACAGTTACCTTACCCCAAGGATGATGTGTCTGGGAGCTTCTGTG 1500
Db 1441 CTACTGTGTATGAGACAGTTACCTTACCCCAAGGATGATGTGTCTGGGAGCTTCTGTG 1500
Qy 1501 GGGCAGAGACTCTCCAGGAGAGACAGCGGGGCTCTTGTCTGTGAGCAAAACAAACG 1560
Db 1501 GGGCAGAGACTCTCCAGGAGAGACAGCGGGGCTCTTGTCTGTGAGCAAAACAAACG 1560
Qy 1561 CTGCTACCTGGCAGAGTGTCAACAGCTGGGGCAACAGCTGTGGCCAGAGAAACAAACCTGG 1620
Db 1561 CTGCTACCTGGCAGAGTGTCAACAGCTGGGGCAACAGCTGTGGCCAGAGAAACAAACCTGG 1620
Qy 1621 TGTGTACCAACCAAGTACAGAAAGTTCTTCCCTGGATTTACAGCAAGATGAGAGAGAGT 1680
Db 1621 TGTGTACCAACCAAGTACAGAAAGTTCTTCCCTGGATTTACAGCAAGATGAGAGAGAGT 1680
Qy 1681 GCGATTGAGAAATCTTAACCAAGCTGGCTGTGCTGTGCAACAGACCGGCTGTGTGAC 1740
Db 1681 GCGATTGAGAAATCTTAACCAAGCTGGCTGTGCTGTGCAACAGACCGGCTGTGTGAC 1740
Qy 1741 TCGAGAAA 1748
Db 1741 TCGAGAAA 1748

RESULT 2
US-10-177-661-1
; Sequence 1, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Virchow, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299,606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
; OTHER INFORMATION:
US-10-177-661-1

Query Match 72.0%; Score 1258.6; DB 3; Length 1434;
Best Local Similarity 98.5%; Pred. No. 2.2e-237;
Matches 1270; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 411 CAGCAGCAGAGCAACGAGGCGCACGAGGAGAGCCAGGATACAGAGCTGCCAAGTTCACT 470
Db 146 CAGCCTATATATTCCTGCTGCGCGCCCTGCTCTAGATGAGAGCTGCCAAGTTCACT 205
Qy 471 GCGCGAGAGGCGCAAGACAGCTACCGCTCATCGGATGAGTCTCTCTCTCAATTCCTGCG 530
Db 206 GCGCGAGAGGCGCAAGACAGCTACCGCTCATCGGATGAGTCTCTCTCTCAATTCCTGCG 265
Qy 531 TGGTTTGGCTCATATCTCTTCCAGTTTGGAGGGGCGACAGGGATCAGGTACAAAG 590
Db TGGTTTGGCTCATATCTCTTCCAGTTTGGAGGGGCGACAGGGATCAGGTACAAAG 590

Db 266 TGGTTGCTCATCTCCTTCCAGTTCTGAGGAGGCAACAGGATCAGTACAGG 325
 QY 591 AGCAGAGGAGAGCTGTCTCCCAAGCAGCTGTTCCTGTGACGGGTGTGACTGCAAGC 650
 Db 326 AGCAGAGGAGAGCTGTCTCCCAAGCAGCTGTTCCTGTGACGGGTGTGACTGCAAGC 385
 QY 651 TGAAGAGTGAAGAGCTGGGCTGCTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 710
 Db 386 TGAAGAGTGAAGAGCTGGGCTGCTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 445
 QY 711 ACTCTGGGTCTCTCCCATCAGTGAGTCTCCCATCTGTAGCAGCACTGGAATGACTCTACT 770
 Db 446 ACTCTGGGTCTCTCCCATCAGTGAGTCTCCCATCTGTAGCAGCACTGGAATGACTCTACT 505
 QY 771 CAGAGAAAGACTGCTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACCAAGGTTGCC 830
 Db 506 CAGAGAAAGACTGCTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACCAAGGTTGCC 565
 QY 831 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCC 890
 Db 566 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCC 625
 QY 891 TCACAGAGTGAATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTTCACACTGCGGAC 950
 Db 626 TCACAGAGTGAATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTTCACACTGCGGAC 685
 QY 951 TGAAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTT 1010
 Db 686 TGAAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTT 745
 QY 1011 GGCAGAGTGAATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTTCACACTGCGGAC 1070
 Db 746 GGCAGAGTGAATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTTCACACTGCGGAC 805
 QY 1071 AGTGGGTGCTCACTGCGGCGCATCTGCTTCTGTGACCCGCGAGAGAGTCTCTGAGGAGCT 1130
 Db 806 AGTGGGTGCTCACTGCGGCGCATCTGCTTCTGTGACCCGCGAGAGAGTCTCTGAGGAGCT 865
 QY 1131 GGAAGGTGATCGGCGGCAACAGCAACTGTGACAGAGTGTCTGAGGAGAGCTCTCATTTGCGG 1190
 Db 866 GGAAGGTGATCGGCGGCAACAGCAACTGTGACAGAGTGTCTGAGGAGAGCTCTCATTTGCGG 925
 QY 1191 AGATCATCATCAAGCAATTAACAACGATGAGAGAGAGCACTATGACATGCGCCCTCATGCG 1250
 Db 926 AGATCATCATCAAGCAATTAACAACGATGAGAGAGAGCACTATGACATGCGCCCTCATGCG 985
 QY 1251 GCGTGTCAAGGCGCCTGACCTGTCCGCTCAATCAACCTGCTTGTGCTCCCATGATG 1310
 Db 986 GCGTGTCAAGGCGCCTGACCTGTCCGCTCAATCAACCTGCTTGTGCTCCCATGATG 1045
 QY 1311 GACAGACCTTTAGCTCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGAG 1370
 Db 1046 GACAGACCTTTAGCTCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGAG 1105
 QY 1371 CAGATGACAAAGATCCCTTCTCCGCGAGGTGAGGTCATCTCATGACTTCAAGA 1430
 Db 1106 CAGATGACAAAGATCCCTTCTCCGCGAGGTGAGGTCATCTCATGACTTCAAGA 1165
 QY 1431 AATCAATGATCTACTTGTCTTATGACAGTTACTTACCCCAAGATGATGTGTGTGGG 1490
 Db 1166 AATCAATGATCTACTTGTCTTATGACAGTTACTTACCCCAAGATGATGTGTGTGGG 1225
 QY 1491 ACCTTGTGGGGGCAAGACTTCTGCGCAGGAGACAGCGGGGGGCTTTGTCTGTGAGC 1550
 Db 1226 ACCTTGTGGGGGCAAGACTTCTGCGCAGGAGACAGCGGGGGGCTTTGTCTGTGAGC 1285
 QY 1551 AGAACAACCGGTGATCTGCGAGGTGACAGAGTGGGGCAAGGCTGTGGCCAGAGAA 1610
 Db 1286 AGAACAACCGGTGATCTGCGAGGTGACAGAGTGGGGCAAGGCTGTGGCCAGAGAA 1345
 QY 1611 ACAAACTGTGTGTACACAAAGTGAAGAGTCTTCTCTGATTTTACAGCAAGATGG 1670
 Db 1346 ACAAACTGTGTGTACACAAAGTGAAGAGTCTTCTCTGATTTTACAGCAAGATGG 1405

QY 1671 AGAGCGAGGTGCGATTCAGAAAAATCTTAA 1699
 Db 1406 AGAGCGAGGTGCGATTCAGAAAAATCTTAA 1434
 RESULT 3
 US-10-177-661-3
 ; Sequence 3, Application US/10177661
 ; Patent No. 6794173
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Dirk M.
 ; APPLICANT: Virca, G. Duke
 ; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
 ; FILE REFERENCE: 3256-A
 ; CURRENT APPLICATION NUMBER: US/10/177,661
 ; PRIOR FILING DATE: 2002-06-20
 ; PRIOR FILING DATE: 2001-06-20
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1341
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1341)
 ; OTHER INFORMATION:
 ; US-10-177-661-3
 Query Match 60.8%; Score 1062.6; DB 3; Length 1341;
 Best Local Similarity 91.3%; Pred. No. 1.5e-249;
 Matches 1177; Conservative 0; Mismatches 19; Indels 93; Gaps 1;
 QY 411 CAGCAGCAGCAACCAAGGCGCAACAGGAGAGCCAGGATAGAGCTGCCAAGTTCACT 470
 Db 146 CAGCCTATATATCCCTGCTGCGGCGCCTGCTCTAGGATAGAGCTGCCAAGTTCACT 205
 QY 471 GCGGAGGCGCAGAGCAGCTACCGCTCATCGGCTGCTCTCTCTCATTTGCCCTGG 530
 Db 206 GCGGAGGCGCAGAGCAGCTACCGCTCATCGGCTGCTCTCTCTCATTTGCCCTGG 265
 QY 531 TGGTTGGCTCATCTCTCTCCAGTTCTGGGAGGCGCACAGGAGTACAGTACAAG 590
 Db 266 TGGTTGGCTCATCTCTCTCCAGTTCTGGGAGGCGCACAGGAGTACAGTACAAG 325
 QY 591 AGCAGAGGAGAGCTGTCCCAAGCAGCTGTTCCTGTGACGGGTGTGACTGCAAGC 650
 Db 326 AGCAGAGGAGAGCTGTCCCAAGCAGCTGTTCCTGTGACGGGTGTGACTGCAAGC 385
 QY 651 TGAAGAGTGAAGAGCTGGGCTGCTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 710
 Db 386 TGAAGAGTGAAGAGCTGGGCTGCTGAGGTTGACTGGGCAAGTCTCTGCTTAAATCT 445
 QY 711 ACTCTGGGTCTCTCCCATCAGTGAGTCTCCCATCTGTAGCAGCACTGGAATGACTCTACT 770
 Db 446 ACTCTGGGTCTCTCCCATCAGTGAGTCTCCCATCTGTAGCAGCACTGGAATGACTCTACT 505
 QY 771 CAGAGAAAGACTGCTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACCAAGGTTGCC 830
 Db 506 CAGAGAAAGACTGCTGACAGAGTGGGTTTCCAGAGTGTCTACCCGCAACCAAGGTTGCC 537
 QY 831 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCC 890
 Db 538 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCC 537
 QY 891 TCACAGAGTGAATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTTCACACTGCGGAC 950
 Db 538 ACAGAGATTTTGGCAAGCTTCTCAATCTTGAGATACAACTCCACCATCCAGGAAAGCC 537
 QY 951 TGAAGGCGCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTT 1010

```
Db      593  TGAGGGCCATGACCGGGCGATCGTGGAGGGGCGCTGGCTCGATGACAAAGTGCCCTT 652
Qy      1011  GGCAAGTAGTCTGCACTTGGGCAACCACTCTGAGGAGCGCTCATTTGACGCCC 1070
Db      653  GGCAAGTAGTCTGCACTTGGGCAACCACTCTGAGGAGCGCTCATTTGACGCCC 712
Qy      1071  AGTGGGTGCTCACTGCGGCCCACTGCTTCTTCTGAGACCCGGAGAAAGTCTTGAAGGCT 1130
Db      713  AGTGGGTGCTCACTGCGGCCCACTGCTTCTTCTGAGACCCGGAGAAAGTCTTGAAGGCT 772
Qy      1131  GGAAGGTGTCAGCGGGGCAACGAACTGCAACAGTTCCTGAGGAGACCTTCATTTGCCG 1190
Db      773  GGAAGGTGTCAGCGGGGCAACGAACTGCAACAGTTCCTGAGGAGACCTTCATTTGCCG 832
Qy      1191  AGATCATCATCAACAGCAATTACACCGATGAGAGAGCACTATGACATGCGCCCTCATGC 1250
Db      833  AGATCATCATCAACAGCAATTACACCGATGAGAGAGCACTATGACATGCGCCCTCATGC 892
Qy      1251  GGTGTCCAGGCCCTGACCTGTCCGCTCAATCCACCTGCTTGCCTCCCATGATG 1310
Db      893  GGTGTCCAGGCCCTGACCTGTCCGCTCAATCCACCTGCTTGCCTCCCATGATG 952
Qy      1311  GACGACCTTTAGCTGAAATGAGACCTGAGATCAACAGCTTTGGCAAGACAGGAGA 1370
Db      953  GACGACCTTTAGCTGAAATGAGACCTGAGATCAACAGCTTTGGCAAGACAGGAGA 1012
Qy      1371  CAGATGACAAAGACATCCCTTCTCCGGAGGTGACAGTCAATCTCATGACTTCAAGA 1430
Db      1013  CAGATGACAAAGACATCCCTTCTCCGGAGGTGACAGTCAATCTCATGACTTCAAGA 1072
Qy      1431  AATGCAATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGG 1490
Db      1073  AATGCAATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGG 1132
Qy      1491  ACCTTCGTGGGGGAGAGACCTTCCTCCAGGAGAGACAGGGGGGGCTTTGTCTGTAGC 1550
Db      1133  ACCTTCGTGGGGGAGAGACCTTCCTCCAGGAGAGACAGGGGGGGCTTTGTCTGTAGC 1192
Qy      1551  AGAACAACCGCTGTGTAACCTGACAGGTGTCACACAGCTGGGAGCAAGGCTGTGGCAGAA 1610
Db      1193  AGAACAACCGCTGTGTAACCTGACAGGTGTCACACAGCTGGGAGCAAGGCTGTGGCAGAA 1252
Qy      1611  ACAAACTGTGTGTACACCAAGTGAAGAAGTTCTTCCCTGTGATTTTACAGCAAGATGG 1670
Db      1253  ACAAACTGTGTGTACACCAAGTGAAGAAGTTCTTCCCTGTGATTTTACAGCAAGATGG 1312
Qy      1671  AGAGCGAGTCCGATTCAGAAAATCTTAA 1699
Db      1313  AGAGCGAGTCCGATTCAGAAAATCTTAA 1341

RESULT 4
US-09-879-792-35
; Sequence 35, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xieo, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
(Docket No. 6734006 L10-81-MO)
```

```
; SEQ ID NO 35
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) .. (1230)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-792-35

Query Match      60.2%; Score 1053; DB 3; Length 1230;
Beet Local Similarity 92.1%; Pred. No. 3,3e-247;
Matches 1134; Conservative 0; Mismatches 84; Indels 13; Gaps 2;

Qy      521  ATTGCGCTGGTGGTTCGCTCATCATCTCTTCCAGTTCTGGCAGGGCCACAGGATC 580
Db      1  ATGACCCAGCTGTGTGCTTCTTCTTCTTCTAGTCCAGTTCTGGCAGNNCCACAGNNATC 60
Qy      581  AGGTACAAAGAGCAGAGAGAGAGCTGTCCAGACGCTGTGCTGTGAACGGGTGGTG 640
Db      61  AGGTACAAAGAGCAGAGAGAGAGCTGTCCAGACGCTGTGCTGTGAACGGGTGGTG 120
Qy      641  GACTGCAAGCTGAAGATGAGAGCTGGCTGCTGAGTTTGAATGGAGCAAGTCTCTG 700
Db      121  GACTGCAAGCTGAAGATGAGAGCTGGCTGCTGAGTTTGAATGGAGCAAGTCTCTG 180
Qy      701  CTTAAATCTACTGTGGGTCTCTCCATCAGTGGCTTCCATCTGTAGCAGCACTGGAAT 760
Db      181  CTTAAATCTACTGTGGGTCTCTCCATCAGTGGCTTCCATCTGTAGCAGCACTGGAAT 240
Qy      761  GACTCTACTCAGAGAAAGACCTGCAAGCTGAGCTGGGTTTGAAGTGTCTACCGGAAAC 820
Db      241  GACTCTACTCAGAGAAAGACCTGCAAGCTGAGCTGGGTTTGAAGTGTCTACCGGAAAC 300
Qy      821  GAGGTTCGCCACAGAGATTTTGGCAACAGCTTCTCAATCTTGAATCAATCCACATC 880
Db      301  GAGGTTCGCCACAGAGATTTTGGCAACAGCTTCTCAATCTTGAATCAATCCACATC 360
Qy      881  CAGGAAGAGCTTCCACAGAGTGTGAATGCTTCCAGAGGATATATCTCTCCAGATGTTCC 940
Db      361  CAGGAAGAGCTTCCACAGAGTGTGAATGCTTCCAGAGGATATATCTCTCCAGATGTTCC 420
Qy      941  CACTGCGAGCTGAGAGGACATGACCGGGCGAGTGTGGAGAGGGGCTGGCTCGGATAGC 1000
Db      421  CACTGCGAGCTGAGAGGACATGACCGGGCGAGTGTGGAGAGGGGCTGGCTCGGATAGC 480
Qy      1001  AAGTGGCTTGGCAAGTGAATGTGCACTTGGGACCAACCAATCTTGTGAAGGACGCTC 1060
Db      481  AAGTGGCTTGGCAAGTGAATGTGCACTTGGGACCAACCAATCTTGTGAAGGACGCTC 540
Qy      1061  ATTGACGCCAGTGGGTGTCTCATGCGGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTTC 1120
Db      541  ATTGACGCCAGTGGGTGTCTCATGCGGCCCACTGCTTCTTCTGTGACCCGGAGAAAGTTC 600
Qy      1121  CTGAGAGGCTGGAAGGTGTACGCGGGGCAACGAACTTGCACAGTTCCTGAGGACGCC 1180
Db      601  CTGAGAGGCTGGAAGGTGTACGCGGGGCAACGAACTTGCACAGTTCCTGAGGACGCC 660
Qy      1181  TCCATTGCCAGATCATCATCAACAGCAATTTACCCGATGAGAGGAGCGACTATGATCATC 1240
Db      661  TCCATTGCCAGATCATCATCAACAGCAATTTACCCGATGAGAGGAGCGACTATGATCATC 720
Qy      1241  GCCCTCATGCGGCTGTCCAGACCCCTGAGCCCTGCTCATCATCAACCTGCTTGGCTTC 1300
Db      721  GCCCTCATGCGGCTGTCCAGACCCCTGAGCCCTGCTCATCATCAACCTGCTTGGCTTC 780
Qy      1301  CCCATGATGAGACAGACCTTTAGCTTCAATGAGACCTGCTGTGATCAAGGCTTTGGCAA- 1359
Db      781  TCTCTGCCCCCGCCAGCCCAACCTCTGACGCTTCCGACTGTGACAGATCTGTCAAC 840
Qy      1360  -----GACGAGGAGACAGAGAGCAAGATCCCTTCCCTCCGGAGGTGCAG 1408
Db      841  TCATATCCGGGCCCCCAAGCTTCTGACAGACAGACATCCCTTCTCCGGAGGTGCAG 900
```


QY	1409	GTCAATCTCATCGACCTTCAGAAAATGCATATACACTTGGTGTATGACAGTTACCTTACC	1468
Db	901	GTCAATCTCATCGACCTTCAGAAAATGCATATACACTTGGTGTATGACAGTTACCTTACC	960
QY	1469	CCAAGGATGATGTGTGCTGGGGGACCTTCGTGTGGGGGACGAGACCTGCTCCAGGGGAGACAGC	1528
Db	961	CCAAGGATGATGTGTGCTGGGGGACCTTCGTGTGGGGGACGAGACCTGCTCCAGGGGAGACAGC	1020
QY	1529	GGGGGGGCTCTTGTCTGTGAGCAGAACAAACCGCTGTACCTGGCAGGTTGTCACCAAGCTGG	1588
Db	1021	GGGGGGGCTCTTGTCTGTGAGCAGAACAAACCGCTGTACCTGGCAGGTTGTCACCAAGCTGG	1080
QY	1589	GGCACAAGCTGTGGCCAGAGAAACAAACCTGTGTGTACACCAATGTACAGAAATCTTT	1648
Db	1081	GGCACAAGCTGTGGCCAGAGAAACAAACCTGTGTGTACACCAATGTACAGAAATCTTT	1140
QY	1649	CCCTTGGATTTTACAGCAAGATGTGAAGAGCAGGTGGATTACAGAAAATCTTAACCAAGCTGGC	1708
Db	1141	CCCTTGGATTTTACAGCAAGATGTGAG- GCCAGGTGTGGATTACAGAAAATCTTAACCAAGCTGGC	1199
QY	1709	CTGCTGCTCTGACAGCAACGGGCTGTGTGA	1739
Db	1200	CTGCTGCTCTGACAGCAACGGGCTGTGTGA	1230

[illegible]

QY	904	ATGCCCTTCCACGGGGTATATCTTCCCTCCAGTGTTCACATCGGGAATTGAGGGGCATATGAC	963
Db	241	ATGCCCTTCCACGGGGTATATCTTCCCTCCAGTGTTCACATCGGGAATTGAGGGGCATATGAC	300
QY	964	CGGCGGATCGTGGGAGGGGGCGTGGGCGCTGGATAGCAAGTGGCTTGGCAAGTAGTCT	1023
Db	301	CGGCGGATCGTGGGAGGGGGCGTGGGCGCTGGATAGCAAGTGGCTTGGCAAGTAGTCT	360
QY	1024	GCACCTTGGGCACCAACCAATCTGTGGAGGCACGCTCATTTGACGCCACAGTGGGTCTCAC	1083
Db	361	GCACCTTGGGCACCAACCAATCTGTGGAGGCACGCTCATTTGACGCCACAGTGGGTCTCAC	420
QY	1084	TGCGCGCCCATCTGCTTCTTCTTGATACCCGGGAGAAAGTCTTGGAGGGCTGGAAAGGTATACGC	1143
Db	421	TGCGCGCCCATCTGCTTCTTCTTGATACCCGGGAGAAAGTCTTGGAGGGCTGGAAAGGTATACGC	480
QY	1144	GGGCAACGACCACTTGCACCGATGGCTGTGAGGCAAGCTCCATTTGCGAGATCATCATCAAA	1203
Db	481	GGGCAACGACCACTTGCACCGATGGCTGTGAGGCAAGCTCCATTTGCGAGATCATCATCAAA	538
QY	1204	CAGCAATTTACACCGATGAGGAGGAGCACTA-TGACATTCGCTTCATGCGGCTGTCCAAAGC	1262
Db	539	CAGCAATTTACACCGATGAGGAGGAGCACTAATGATATCGCTTCATGCGGCTGTTCCAAAG	598
QY	1263	CCCTGACCCCTGTCCGCTTCATCATCCACCTGTGTCCCT-CCCCATGCAATGACAAACCTTT	1321
Db	599	CCCTGACCACTGTCCGCTTCATCATCCACCTGTGTCCCTCCCCCAATGCAATGACAAACCTTT	658
QY	1322	AGCCTCAATGAGACTGTGTCATACAGGCTTTGAGCAAGACAGGAGAGACAGATGACAAAG	1381
Db	659	AGCCTCAATGAGACTGTGTCATACAGGCTTTGAGCAAGACAGGAGAGACAGATGACAAAG	718
QY	1382	ACATTCGCCCTTCTTCGGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGAC	1441
Db	719	ACATTCGCCCTTCTTCGGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGAC	777
QY	1442	TACTTGGTCTATGACAGTACCTTAACCCCAAGATGATGTGTGCTGGGAGCCT	1494
Db	778	TACTTGGTCTATGACAGT-ACCTTACCCCAAGATGATGTGTGCGGGAACTT	828

```

RESULT 6
US-09-879-792-30
; Sequence 30, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonhong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; _____ (Docket No. 6734006 LIO-81-WO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 678
; TYPE: DNA
; ORGANISM: mouse
US-09-879-792-30

Query Match      25.2%; Score 440.8; DB 3; Length 678;
Best Local Similarity 82.7%; Pred.No. 7.4e-98;
Matches 516; Conservative 0; Mismatches 107; Indels 1; Gaps 1;

```

QY 320 TCAGCTGCTGAGCAACCTCCCAACAGAGTGTACTTGTAGAGCAACAGAGTGGG 379
Db 1 TCAGCTGCTGAGCAACCTCCCAACAGAGTGTACTTGTAGAGCAACAGAGTGGG 60
QY 380 GCTGTACCCATCCGATCATCTCTGCGAGGTGAGCAACAGAGTGGGCAACAGAGGAG 439
Db 61 GCTGTACCCATCCGAGCATCTCTGCGAGGTGAGCAACAGAGTGGGCAACAGAGGAG 120
QY 440 AGCCAGGTGAGAGGCTGAGAGTTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 499
Db 121 AGCCAGGTGAGAGTTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 500 ATCGAGTGTGCT 559
Db 181 ATCGAGTGTGCT 240
QY 560 TGGCAGGCGCCACAGAGGATGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 619
Db 241 TGGAGAGGCGCCACAGAGGATGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 620 GTTGCTGTGAGCGGAGTGTGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
Db 301 GTTGCTGTGAGCGGAGTGTGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 680 TTGAGTGGGCAAGTCT 739
Db 361 TTGAGTGGGCAAGTCT 420
QY 740 ATCTGTAGAGCAACTGGAATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 799
Db 421 GTCTGAGCAGCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 800 GAGAGTGTCAACCGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
Db 481 GAGAGTGTCAACCGAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 860 TTGAGTACCACTCCAGCAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919
Db 541 TCCGAGTACCACTCCAGCAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
QY 920 TATATCT 943
Db 600 TATATCT 623

RESULT 7
US-09-879-792-33
Sequence 33, Application US/09879792
Patent No. 6734006
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Seline
FILE OF INVENTION: Procease
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 614
TYPE: DNA
ORGANISM: mouse
US-09-879-792-33

Query Match 23.5%; Score 411.6; DB 3; Length 614;
Best Local Similarity 90.1%; Pred. No. 9.6e-91;
Matches 463; Conservative 0; Mismatches 49; Indels 2; Gaps 2;
QY 1189 CGAGATCATCATCAACAGCAATTAACCCGATGAGAGAGAGAGAGAGAGAGAGAGAG 1248
Db 1 CCAGATCATCATCAACAGCAATTAACCCGATGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 1249 GCGGCTGTCCAGGCGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1308
Db 61 CAGGCTGTCCAGGCGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 120
QY 1309 TGGACAGAGCTTTAGCTTCAATGAGAGCTGAGAGCTGAGAGCTTTGAGAGAGAG 1368
Db 121 CGGTCAAGAGCTTGGGCTCAATGAGAGCTG-TGGATCAGGGGCTTGGCAAAACCAAG 178
QY 1369 GACAGATGACAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1428
Db 179 GACAGATGACAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 238
QY 1429 GAAATGAGATGACTTGGTCTATGAGAGTACCTTACCCAGAGAGAGAGAGAGAG 1488
Db 239 GAAATGAGATGACTTGGTCTATGAGAGTACCTTACCCAGAGAGAGAGAGAGAG 298
QY 1489 GAGCTTCTGAGGAGAGAGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1548
Db 299 GAGCTTCTGAGGAGAGAGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 358
QY 1549 GCAGAGAACCGCTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1608
Db 359 GCAGAGAACCGCTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
QY 1609 AAAAGAACTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1668
Db 419 AAAAGAACTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 478
QY 1669 GAGAGGAGAGGTGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1702
Db 479 GAGAGGAGAGGTGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 512

RESULT 8
US-09-879-792-1
Sequence 1, Application US/09879792
Patent No. 6734006
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Seline
FILE OF INVENTION: Procease
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879,792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 402
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) - (402)
OTHER INFORMATION: n = A,T,C or G
US-09-879-792-1

Query March 19.9%; Score 348; DB 3; Length 402;
Best Local Similarity 97.8%; Pred. No. 2.8e-75;
Matches 362; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 903 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCAGTGGAGGCGCATGA 962
DB 1 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCAGTGGAGGCGCATGA 60

QY 963 CCGGCGGATGTGTGGAGGGGGCGCTGCTCGATAGCAAGTGGCTTTGGCAAGTACTC 1022
DB 61 CCGGCGGATGTGTGGAGGGGGCGCTGCTCGATAGCAAGTGGCTTTGGCAAGTACTC 120

QY 1023 TGCACTTGGGACCAACCACTCTGTGAGGACGCTATGACGCCCAAGTGGGCTCA 1082
DB 121 TGCACTTGGGACCAACCACTCTGTGAGGACGCTATGACGCCCAAGTGGGCTCA 180

QY 1083 CTGCGCGCCACTGTCTTCTGTG-AACCGGAGAAAGTCTGTGAGGGCTGGAAGTGTAC 1141
DB 181 CTGCGCGCCACTGTCTTCTGTG-AACCGGAGAAAGTCTGTGAGGGCTGGAAGTGTAC 240

QY 1142 GCGGCGACCAACCACTGTGAGGAGGCTGTGAGGAGCTTCATTCCTGAGATCATTC 1201
DB 241 GCGGCGACCAACCACTGTGAGGAGGCTGTGAGGAGCTTCATTCCTGAGATCATTC 300

QY 1202 AACGCAATTACCGGATGAGAGAGACGATATGACATCGCCCTCATGCGGCTGCCAAG 1261
DB 301 AACGCAATTACCGGATGAGAGAGACGATATGACATCGCCCTCATGCGGCTGCCAAG 360

QY 1262 CCCCTGACC 1271
DB 361 NCCCTTGACC 370

RESULT 9
US-09-879-792-29
; Sequence 29, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedlich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879.792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12 (Docket No. 6734006 LIO-81-WO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 621
; TYPE: DNA
; ORGANISM: mouse
US-09-879-792-29

Query Match 19.2%; Score 335.8; DB 3; Length 621;
Best Local Similarity 80.8%; Pred. No. 3.1e-72;
Matches 489; Conservative 0; Mismatches 107; Indels 9; Gaps 8;

QY 302 AGGTATATCCGAGGATGAGCTCGGTGACACCTCCCAACAGAGTGAACCTTGT 361
DB 1 AGATATATCTGCGAGGATGAGCTCGGTGACACCTCCCAACAGAGTGAACCTTGT 60

QY 362 AGAGCAACACGAGTGGGGCTGTACCATTCGATCATCTCTGCGAGGTGAGACACAGA 421
DB 61 AGAGCAACACGAGTGGGGCTGTACCATTCGATCATCTCTGCGAGGTGAGACACAGA 120

QY 422 ACCAGGGCCACAGAGG-AGAGCCAGGATACGAGGCTGCCCAAGTTCACTGG-CGGAGG 479
DB 121 ACCAGGGCCACAGAGGATGAGCCAGGCTTCAAGTTTCCCAAGTTCTCTGTGAGAGA 180

QY 480 GCCAAGACAGCTACCGCTCATCGGATGCGGTCTCTCTCATTTGCCGTGGTGTTCGC 539
DB 181 CCCAAGACAGCTACCGCTCATCGGATGCGGTCTCTCTCATTTGCCGTGGTGTTCGC 240

QY 540 TCATCATCTCTTCCAGTTCGG-CAGGGCCACACAGGATCAGATCAAGAGCAGAGG 598
DB 241 TCATCATCTCTTCCAGTTCGG-CAGGGCCACACAGGATCAGATCAAGAGCAGAGG 300

QY 599 GAGAGCTGTCCCAAGACGCGTTCGCTGAGCGGGGTGGAGC-TGCAAGTGAAGAG 657
DB 301 GAGAGCTGTCCCAAGACGCGTTCGCTGAGCGGGGTGGAGC-TGCAAGTGAAGAG 360

QY 658 TGACAGCTGGGCTGCGTGAAGTGTGACTGAGACAGTCTCTGTAATACTACTCTGG 717
DB 361 CGATGAGCTGGGCTGCGTGAAGTGTGACTGAGACAGTCTCTGTAATACTACTCTGG 420

QY 718 GTCTTC-CCATCATGCTTCCATCTGTGACAGCACTGAAATGATCTCTACTAGGA 776
DB 421 GTCTTCGCGAGAGTGGCTTCTGTCTGTGACAGCAGC-GGAAAGCACTGACTCAGGA 479

QY 777 AGACCTGCGAGC-AGCTGGGTTGAGAGTGTCTACCGGACACCGAGGTTGCCACAGG 835
DB 480 GAGCTGCGAGCAGCTGGAGTTGACAGCTTACCGAACAATGAGGTAGCCACAGA 539

QY 836 GATTTGCGAAGCAGCTTCTCATCTTGAATACAACTCCACATCCAGAGAAAGCTCCAG 895
DB 540 GATTTGCGAAGCAGCTTCTCATCTTGAATACAACTCCACATCCAGAGAAAGCTCCAG 597

QY 896 AGGTC 900
DB 598 AGGTC 602

RESULT 10
US-09-280-116-22
; Sequence 22, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280.116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: trypsin-like serine proteases
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(610)
; OTHER INFORMATION: n = a, t, c, or g
US-09-280-116-22

Query Match 19.1%; Score 333.2; DB 3; Length 610;
Best Local Similarity 97.0%; Pred. No. 1.3e-71;
Matches 360; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 903 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCAGTGGAGGCGCATGA 962
DB 1 AATGCCCTTCCAGCGGATATCTCCCTCCAGTGTTCCTCCAGTGGAGGCGCATGA 60

QY 963 CCGGCGGATGTGTGGAGGGGGCGCTGCTCGATAGCAAGTGGCTTTGGCAAGTACTC 1022
DB 61 CCGGCGGATGTGTGGAGGGGGCGCTGCTCGATAGCAAGTGGCTTTGGCAAGTACTC 120

```

QY      1023  TGCACCTTGGGACGACCCCACTCTGTGAGGACGCTCATTTGACGCCCGAGTGGTGTCA 1082
Db      121  TGCACCTTGGGACGACCCCACTCTGTGAGGACGCTCATTTGACGCCCGAGTGGTGTCA 180
QY      1083  CTGCGCCGACCTGCTTTCTTCTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAGGTGTACG 1142
Db      181  TTGGGCGCCGACCTGCTTTCTTCTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAGGTGTACG 240
QY      1143  CGGGCA-CCAGCAACCTGCA-CCAGTTGCTGAGGACGCTCCATTGCGGAGATCATCAT 1200
Db      241  CGGGCAACCCAGAACCTGCAACCCAGTTGCTGAGGACGCTCCATTGCGGAGATCATCAT 300
QY      1201  CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGATGACATCGCCCTCATGCGGCTGTCCA 1260
Db      301  CAACAGCAATTACACCGATGAGAGAGAGAGAGAGAGATGACATCGCCCTCATGCGGCTGTCCA 360
QY      1261  GCCCTTGACC 1271
Db      361  GCCCTTGACC 371

```

RESULT 11

```

US-09-879-792-2
; Sequence 2, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-792-2

```

```

Query Match      16.3%; Score 285; DB 3; Length 285;
Best Local Similarity 100.0%; Pred. No. 5.8e-60;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      957  CCATGACCGGCGGATCTGTGAGGAGGCGCTGCGCTCGATAGCAAGTGCGCTTGGCAAG 1016
Db      1  CCATGACCGGCGGATCTGTGAGGAGGCGCTGCGCTCGATAGCAAGTGCGCTTGGCAAG 60
QY      1017  TGAATCTGCACTTGTGGGACCAACCACTGTGAGAGGACGCTCATTTGAGCCCAAGTGG 1076
Db      61  TGAATCTGCACTTGTGGGACCAACCACTGTGAGAGGACGCTCATTTGAGCCCAAGTGG 120
QY      1077  TGCTCACTGCGCCCACTGCTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAG 1136
Db      121  TGCTCACTGCGCCCACTGCTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAG 180
QY      1137  TGTACGCGGAGACCAAGCACTGTGACCAAGTTGCTTGAAGAGCCTTCATTGCCAGATCA 1196
Db      181  TGTACGCGGAGACCAAGCACTGTGACCAAGTTGCTTGAAGAGCCTTCATTGCCAGATCA 240
QY      1197  TCATCAACAGCAATTACCCGATGAGAGAGAGAGAGACTATGACATCG 1241
Db      241  TCATCAACAGCAATTACCCGATGAGAGAGAGAGAGACTATGACATCG 285

```

RESULT 12

```

US-09-879-792-5/c
; Sequence 5, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: US/09/879, 792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-792-5

```

```

Query Match      15.8%; Score 276.4; DB 3; Length 286;
Best Local Similarity 97.9%; Pred. No. 7.3e-58;
Matches 280; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      957  CCATGACCGGCGGATCTGTGAGGAGGCGCTGCGCTCGATAGCAAGTGCGCTTGGCAAG 1016
Db      286  CCATGACCGGCGGATCTGTGAGGAGGCGCTGCGCTCGATAGCAAGTGCGCTTGGCAAG 227
QY      1017  TGAATCTGCACTTGTGGGACCAACCACTGTGAGAGGACGCTCATTTGAGCCCAAGTGG 1076
Db      226  TGAATCTGCACTTGTGGGACCAACCACTGTGAGAGGACGCTCATTTGAGCCCAAGTGG 167
QY      1077  TGCTCACTGCGCCCACTGCTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAG 1136
Db      166  TGCTCACTGCGCCCACTGCTTCTTGTGACCCCGGAGAGAGTCTCTGAGGGCTGGAAG 107
QY      1137  TGTACGCGGAGACCAAGCACTGTGACCAAGTTGCTTGAAGAGCCTTCATTGCCAGATCA 1196
Db      106  TGTACGCGGAGACCAAGCACTGTGACCAAGTTGCTTGAAGAGCCTTCATTGCCAGATCA 47
QY      1197  TCATCAACAGCAATTACCCGATGAGAGAGAGAGAGACTATGACATCG 1242
Db      46  TCATCAACAGCAATTACCCGATGAGAGAGAGAGAGACTATGACATCG 1

```

RESULT 13

```

US-09-879-792-34
; Sequence 34, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: US/09/879, 792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16

```

PRIOR APPLICATION NUMBER: PCT _____ (Docket No. 6734006 LIO-81-WO)
 PRIOR FILING DATE: 2001-06-12
 NUMBER OF SEQ ID NOS: 36
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 34
 LENGTH: 751
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-879-792-34

Query Match 15.2%; Score 265.4; DB 3; Length 751;
 Best Local Similarity 73.3%; Pred. No. 5.1e-55;
 Matches 486; Conservative 0; Mismatches 141; Indels 36; Gaps 10;

183 AGGATCTCCAGCTGTAACCTTCAGGCGG-GCATTCAGGCGCGGATCTCCAGCC 241
 53 ATGATCTCCAGCAAGAACCTTCAGCTGAGCATCTCCAGGCGGATCTCCAGCT 112
 242 CAGGATCTCCAGGCGGCGATCTCCGCTGAGCATCTTCAGGCTCTCATCC-GG 300
 113 GGGACACTCCAGGCGGCGGATCTTCAGCCAGGATCACTTCAGGCTCTCATCC 172
 301 CAGGATCTCCAGGCGGCGATCTCCGCTGAGCAACCTCCCAACAGAGTGAACCT 360
 173 CAGGATCTCCAGGCGGCGATCTTCAGGCTGAGCAACCTCCCAACAGAGTGAAC 232
 361 TAGAGCAACACAGTGGGGCTGTAACCATTCATCTCTTCAGAGTGAACAGC 420
 233 TAGAGCAACACAGTGGGGCTGTAACCATTCATCTCTTCAGAGTGAACAGC 292
 421 AACAGGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 478
 293 AACAGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 352
 479 GGGCAGAGAGAGTACCGCTCATCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 533
 353 GGGCAGAGAGAGTACCGCTCATCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 412
 534 -TTTCGCTCATCATCTCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 585
 413 TTTCGCTCATCATCTCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 472
 586 CAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 644
 473 GCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 532
 645 GCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 689
 533 GACTTCCAAAGAGTGAAGAGTGAAGAGGCTGAGGCTGAGGCTGAGGCTGAGG 592
 690 -ACAAGTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 747
 593 AACGCTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 652
 748 CAGGAGCT-GGAAGTACTCTACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 806
 653 CAGGAGCTGGGAATTGACTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
 807 CTC 809
 713 CCC 715

RESULT 14
 US-09-008-271A-18

Sequence 18, Application US/0908271A
 Patent No. 6203979

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
 Hillman, Jennifer L.
 Yue, Henry
 Guegler, Karl J.
 Corley, Neil C.

Tang, Tom Y.
 Shah, Purya
 TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/008,271A

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mohan-Peterson, Sheela

REGISTRATION NUMBER: 41,201

REFERENCE/DOCKET NUMBER: PF-0458 US

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 2038 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNMOT13

CLONE: 1337018

SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Query Match 11.1%; Score 193.2; DB 3; Length 2038;
 Best Local Similarity 55.8%; Pred. No. 3.1e-37;
 Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

924 TCTTCCTCAGTGTCTCCACTGCGGAGTGAAGGCGGATGAGCGGCGGATGCTGGAGGG 983
 759 TCTTCCTCAGTGTCTCCACTGCGGAGTGAAGGCGGATGAGCGGCGGATGCTGGAGGG 818
 984 CGGCGGCTCGGAGTGAAGGCGGCTTGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1043
 819 AGGAGGCTCTGAGATCTTGGCTTGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 878
 1044 TCTGTGAGGAGCAGCTATGAGCGCCAGAGTGGTGTCACTGCCGCCACTGCTTCTTGG 1103
 879 TCTGTGAGGAGGAGATCTTGAAGCCCACTGGGCTTCAAGGAGGCCCACTGCTTCTTGG 934
 1104 TGACCGGAGAGAGTCTTGAAGGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1163
 935 -AGGAACAATACGATGTGTTCAACTGGAAGTGGCGGAGGCTCAGACAACTGGGCA 992
 1164 AGTTCCTGAGGAGCGCTTCATTTGGCGAGATCATCATCAACAGATTAACCGATGAGG 1223
 993 GCTTCC--ATCCCTGAGTGGCGAAGATCATCATTAATTAACCCCATGATCC 1049
 1224 AGGAGCATATGAGCATGCGCTCATGCGGCTGTCCAGGCGGCTGACCTGTCCGCTACA 1283
 1050 CCAAGAGCATATGAGCATGCGCTCATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1109
 1284 TCAACCTGCTGCTTCCCATGATGAGAGAGAGCTTAAAGCTTAAGAGAGAGAGAGAGAG 1343
 1110 TCAGGCCATCTGCTGCTTCTTGTATGAGAGAGCTCATTCAGGAGCCCATCTTGA 1169

QY	1344	TCACAGGCTTTGGCAAGACCAAGGGAGACAGATGACAAGACATCCCCCTCTCCGGAGAG	1403
Db	1170	TCATTGGATGGGGCTTTACGAGACGAATGAGGGAGAGATGTCTGACACTACTGCTCGAGG	1229
QY	1404	TGCAGGTCAATCTCATCGACTTTCAAGAAATGCATGACTTACTTGCTATGACAGTTACC	1463
Db	1230	CGTAGTTCAGGTCAATTGACAGCACACGGTGCATTGACAGCATGCGTACCAAGGGGGAG	1289
QY	1464	TTACCCCCAAGAGTATGTGTCTGGGGACCTTGGTGGGGCAGAGACTCTGCCAGGGAG	1523
Db	1290	TCACCGAAGAAATGATGTGTGTCAGGACATCCCGAAGGGGGTGTGGACACCTGCCAGGGTG	1349
QY	1524	ACAGCGGGGGGCTCTTGTCTGTGAGCAGAAACAACCGTGTACTGTGCAGGTGTCAACA	1583
Db	1350	ACAGTGTGGGGCCCCCTGATGT---ACCAATCTGACCAAGTGCAGTGTGGGCATCGTTA	1406
QY	1584	GCTGGGGCACAGGCGTGGCCAGAGAAACAACCTGGGTGTATACCAAAATGACAGAG	1643
Db	1407	GCTGGGGGTATGGCTGCGGGGGCCCCGAGACCCAGAGAGTATPACCAACAGGTCTTCAGCTT	1466
QY	1644	TTCTTCCCTGGATTTCACGAAGAATGAGACGAGTGCATTCAGAAATCTCAACAG	1703
Db	1467	ATCTCACTGATCTACCAATGTCTGGAAAGGTGAGCTGTAAATGCTGCTGCCCCCTTTGAG	1526

RESULT 15
US-09-968-415-18

```

; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
;
; NUMBER OF SEQUENCES: 24
;
; CORRESPONDENCE ADDRESS:

```

COMPUTER READABLE FORM:

SEQUENCE DESCRIPTION:	SEQ ID NO:
US-09-968-415-18	18

Query Match	11.1%;	Score 193.2;	DB 3;	Length 2038;
Best Local Similarity	55.84;	Pred. No. 3.1e-37;		
Matches 435; Conservative	0;	Mismatches 333;	Indels 12;	Gaps 3

OY	924	TCCTCCCTCAGATGTTTCCCATCTGACGGACCTGAGAGGCGCATGACCGAGGCGCATCGTGGAGAGGGG	983
Db	759	TCTCCCTGACATGCTCTTGCCCTGTGTGGGAGAGCCCTGAAGAGACCCCCGTTGTGGTGGGTGGG	818
OY	984	CGCTGGCCTCGGATAGCAAGTGGCTTTGGCAAGTGAATCTGCACTTGGGACCAACCA	1043
Db	819	AGGAGGCGCTGTGGAATTTCTTGGCTTGGCAGGTCAGCATCTCAAGTAGACAAACAGACAG	878
OY	1044	TCCTGAGAGGCAAGCTCATTTGACGCCCCAGTGGGTCTCATCTGCCCCCATCTGCTTTCTTGG	1103
Db	879	TCTGTGAGAGGAGGATCTTGAGACCCCACTGGGGTCTTCACGGACGCCCATCTGCTTCC---	934
OY	1104	TGACCCGGGAGAGGTCCTGAGAGGGCTGGAAGGTGATGCGGGGACACGACCACTTGACCC	1163
Db	935	--AGGAAACATACCGATGTGTTCACACTGGAAGGTGGGGGACGGCTCAGACAAATGGGGCA	992
OY	1164	AGTTGCGCTGAGGCGACGCTCCATTGGCGAGATTCATCATCAACAGCAATTAACCGGATAGG	1223
Db	993	GCTTCCC---ATCCCGTGGCTGTGGCCAAAGATCATCATCATTTGAATTCACCCCATGTATCC	1049
OY	1224	AGGACGACTATGACATATGGCCCTCATTGGCGCTGTCCAAAGCCCTTGACCCCTGCGGCTCACA	1283
Db	1050	CGAAAGCAATGACATGATGCGCTTCATGGAAGCTGACGTTCCCACTCATCTTTCTTCAGGGCACAG	1109
OY	1284	TTCACCGCTGCTTGGCTCCCCATGATAGACAGACCTTTAGCCTCAATGAGACCTGTGGA	1343
Db	1110	TCAGGCCCATCTGTCTGTGCCCTTTCTTTATAGAGAGCTCATCTCCAGCCACCCCATCTTGG	1169
OY	1344	TCACAGGCTTTGGCAGAACCAAGGAGACAGATGACAAAGACATCCCTTCTCTCCGGAGG	1403
Db	1170	TCATTTGATGGGGCTTTACGAAGCAGATGGAAGGGAAGATGTCTGACATCTGTCTGCAGG	1229
OY	1404	TGCAAGTTCATCTCATGACTTTCAGAAATGCAATGACTACTTGTGTCTATGACAGTTACC	1463
Db	1230	CGTCAGTTCAGGTCATTGACAGCACACGGTGTCAATGACAGATGCGTATCCAGGGGGAG	1289
OY	1464	TTTACCCCAAGGATGATGTGTCTGGGGACCTGTGGGGGACAGACCTCTGCCAGGGAG	1523
Db	1290	TCACCGAAGAAATATATGTGTGACAGGCATCTCCGGAAGGGGGTGTGACACCTGCCAGGGT	1349
OY	1524	ACAGCGGGGGGCTCTTGTCTGTGAGCAGAACCAACGCTGTATCTGGCAGGTGTCAACA	1583
Db	1350	ACATGTGTGGGCCCTGTATGT---ACCAATCTGACAGTGGCATGTGTGGTGGGATCTGTTA	1406
OY	1584	GCTGGGGGACAGGCTGTGGCCAGAGAAACAAACTGTGTGTATACCCAAATGACAGAG	1643
Db	1407	GCTGGGGCTATATGCTGGGGGGGCCCGAGCACCCGAGAGATTAACACCAAGGTCTCAAGCTT	1466
OY	1644	TTCTTCCCTGGATTTACGACAAAGATGAGAGCGAGGTGCGATTCAGAAATACCTAACAG	1703
Db	1467	ATCTCAATGGATCTTACAAATGTCTGGAAGGCTGAGCTGTATGTCTGCTGCCCTTTTGCAG	1526

Search completed: September 17, 2006, 16:00:07
Job time : 571 secs

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New). Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

Result					Description	
# No.	Score	Query Match	Length	ID		
1	1748	100.0	1748	3	US-09-879-792-11	Sequence 11, Appl
2	1748	100.0	1748	9	US-10-806-370-11	Sequence 11, Appl
3	1689	96.6	1707	8	US-10-428-275-129	Sequence 129, Appl
C	1689	96.6	1708	8	US-10-428-275-147	Sequence 147, Appl
	1687.4	96.5	1689	7	US-10-156-214-23	Sequence 23, Appl
	1687.4	96.5	1689	10	US-10-916-758-93	Sequence 93, Appl
	1687.4	96.5	2393	7	US-10-153-6190-99	Sequence 99, Appl
7	1670.4	95.6	2067	7	US-10-156-214-25	Sequence 25, Appl
8	1663.4	95.2	2067	7	US-10-916-758-1	Sequence 1, Appl
9	1663.4	95.2	2067	10	US-10-428-275-103	Sequence 103, Appl
10	1645.4	94.1	2432	8	US-10-428-275-131	Sequence 131, Appl
11	1643.8	94.0	2432	8	US-10-428-275-149	Sequence 149, Appl
12	1643.8	94.0	2432	8	US-10-428-275-141	Sequence 141, Appl
13	1643.8	94.0	2432	8	US-10-428-275-123	Sequence 123, Appl
14	1640	93.8	1771	8	US-10-428-275-119	Sequence 119, Appl
15	1640	93.8	1779	8	US-10-428-275-119	Sequence 119, Appl
16	1479	84.6	2310	8	US-10-428-275-127	Sequence 127, Appl
17	1479	84.6	2324	10	US-10-956-157-2559	Sequence 2559, Appl

18	1429	81.8	1614	3	US-09-886-65-45	Sequence 45, Appl
19	1329	81.8	1614	15	US-11-037-643-45	Sequence 45, Appl
20	1369	78.3	1626	8	US-10-428-275-121	Sequence 12, Appl
21	1302.8	74.5	1314	3	US-09-896-837A-17	Sequence 14, Appl
22	1302.8	74.5	1314	3	US-09-896-837A-17	Sequence 17, Appl
23	1302.8	74.5	1314	8	US-10-428-275-131	Sequence 13, Appl
24	1277.8	73.1	1574	7	US-10-274-659-19	Sequence 29, Appl
25	1277.8	73.1	1574	8	US-10-333-574-29	Sequence 29, Appl
26	1258.6	72.0	1434	6	US-10-177-661-1	Sequence 1, Appl
27	1258.6	72.0	1434	10	US-10-910-507-1	Sequence 1, Appl
28	1144.4	65.5	1146	8	US-10-428-275-133	Sequence 13, Appl
29	1120.4	64.1	1218	8	US-10-428-275-125	Sequence 13, Appl
30	1120.4	64.1	1218	8	US-10-428-275-137	Sequence 13, Appl
31	1105	63.2	1203	8	US-10-428-275-111	Sequence 11, Appl
32	1103.4	63.1	1203	8	US-10-428-275-117	Sequence 11, Appl
33	1101.8	63.0	1203	8	US-10-428-275-115	Sequence 11, Appl
34	1068.4	61.1	1078	3	US-09-898-837A-16	Sequence 16, Appl
35	1068.4	61.1	1341	6	US-10-177-661-3	Sequence 3, Appl
36	1062.6	60.8	1341	10	US-10-910-507-3	Sequence 3, Appl
37	1053	60.2	1230	3	US-09-879-792-35	Sequence 35, Appl
38	1053	60.2	1230	9	US-10-806-370-35	Sequence 35, Appl
39	909	52.0	1130	8	US-10-428-275-113	Sequence 11, Appl
40	906.2	51.8	1101	8	US-10-428-275-109	Sequence 10, Appl
41	782.4	44.8	1222	3	US-09-804-153-5	Sequence 5, Appl
42	782.4	44.8	1222	3	US-09-946-63-3	Sequence 3, Appl
43	782.4	44.8	1222	6	US-10-125-457-3	Sequence 3, Appl
44	782.4	44.8	1222	6	US-10-067-761-5	Sequence 5, Appl
45	782.4	44.8	1222	7	US-10-319-519-5	Sequence 5, Appl

ALIGNMENTS

```

.RESULT 1
US-09-879-792-11
: Sequence 11, Application US/09879792
: Patent No. US20020061850A1
: GENERAL INFORMATION:
: APPLICANT: Xiao, Yonghong
: APPLICANT: Gedrich, Richard
: TITLE OF INVENTION: Regulation of Human Transmembrane Serine
: TITLE OF INVENTION: Protease
: FILE REFERENCE: 02973.00035
: CURRENT APPLICATION NUMBER: US/09/879,792
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/211,224
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: US 60/283,353
: PRIOR FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/283,648
: PRIOR FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: PCT _____
: PRIOR FILING DATE: 2001-06-12
: (Docket No. US20020061850A1 LIO-81-WO)
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11
: LENGTH: 1748
: TYPE: DNA
: ORGANISM: Homo sapiens
:
US-09-879-792-11

```

	Query Match	100.0%	Score 1748;	DB 3;	Length 1748;	
	Best Local Similarity	100.0%	Pred. No. 0;			
	Matches 1748;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Oy	1	CTCAGAGCACTGGAGAGGGAACGCCAGCGGAAATGCATCTCCAGCAAGAACAACCTTCACG	60			
Dd	1	CTCGAGAGCCAATTGAGAGGGGACACGCCAGGGAATGCATCTCCAGCAAGAAAACCTTCACG	60			
Oy	61	TGAGAGCATCTCCAGCCCCAGGCGATCTCCAGCTGGGACACTTCCAGGCCGGGATCTCCAGC	120			
Dd	61	TGGAGCATCTCCAGCCCAGGCGATCTCCAGCTGGGACACTTCCAGGCCGGGATCTCCAGC	120			

Oy 121 CCAGGATCTCCAGGCCAGGAGCTCTCCAGTGGAGACACTCCGGGCGGAGCATCTCCAGC 180
 Db 121 CCAGGATCTCCAGGCCAGGAGCTCTCCAGTGGAGACACTCCGGGCGGAGCATCTCCAGC 180
 Oy 181 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 240
 Db 181 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 240
 Oy 241 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 300
 Db 241 CCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 300
 Oy 301 CAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 360
 Db 301 CAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGGATCTCCAGG 360
 Oy 361 TAGAGCAACACAGTGGGGGCTGTATCCCATCCGATCATCTCTGGCAGGTACAGCACAGC 420
 Db 361 TAGAGCAACACAGTGGGGGCTGTATCCCATCCGATCATCTCTGGCAGGTACAGCACAGC 420
 Oy 421 AACAGGGCCACAGAGGAGGAGCCAGGTACGAGCTGGCCAAAGTTACCTGGCGGAGGG 480
 Db 421 AACAGGGCCACAGAGGAGGAGCCAGGTACGAGCTGGCCAAAGTTACCTGGCGGAGGG 480
 Oy 481 CCAGAGGAGCTACCGCTCATCGAGTGCCTCTCTCTCATTTGCTGTGTGCT 540
 Db 481 CCAGAGGAGCTACCGCTCATCGAGTGCCTCTCTCTCTCATTTGCTGTGTGCT 540
 Oy 541 CATCT 600
 Db 541 CATCT 600
 Oy 601 GAGCTGTCCCAAGCAAGCTGTGTGAGCGGGTGGTGAAGTCAAGCTGAAGAGTGA 660
 Db 601 GAGCTGTCCCAAGCAAGCTGTGTGAGCGGGTGGTGAAGTCAAGCTGAAGAGTGA 660
 Oy 661 GAGCTGTGGGCTGGAGGATTGAGTGGAGCAAGTCTGCTTAAATCTACTCTGGGCTC 720
 Db 661 GAGCTGTGGGCTGGAGGATTGAGTGGAGCAAGTCTGCTTAAATCTACTCTGGGCTC 720
 Oy 721 CTCCCATAGTGGCTTCCCATCTGTAGAGCAAGTGGAAATGCTCTCTCTAGAGAAAG 780
 Db 721 CTCCCATAGTGGCTTCCCATCTGTAGAGCAAGTGGAAATGCTCTCTCTAGAGAAAG 780
 Oy 781 CTGCGAGAGCTGGGTTTGAAGAGTCTCAACGAGCAACGAGGTTGCCACAGGATTT 840
 Db 781 CTGCGAGAGCTGGGTTTGAAGAGTCTCAACGAGCAACGAGGTTGCCACAGGATTT 840
 Oy 841 TGCCAAAGCTTCTCAATCTTGAGATCAACTCCATCCAGAAAGCTTCCACAGGTC 900
 Db 841 TGCCAAAGCTTCTCAATCTTGAGATCAACTCCATCCAGAAAGCTTCCACAGGTC 900
 Oy 901 TGAATGGCTTCCCGAGGATATCTCTCTCCAGTGTCCCAAGGAGCTGAGGAGGAT 960
 Db 901 TGAATGGCTTCCCGAGGATATCTCTCTCCAGTGTCCCAAGGAGCTGAGGAGGAT 960
 Oy 961 GAGCGGGGAGTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 Db 961 GAGCGGGGAGTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 Oy 1021 TCTGCACTTCCGAGCCACCTCATCTGTGAGAGCAAGCTCATTTAGCGCCAGTGGTCT 1080
 Db 1021 TCTGCACTTCCGAGCCACCTCATCTGTGAGAGCAAGCTCATTTAGCGCCAGTGGTCT 1080
 Oy 1081 CACTGCGGCGGAGCT 1140
 Db 1081 CACTGCGGCGGAGCT 1140
 Oy 1141 CGCGGGGAGCAAGCAAGCTGCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 Db 1141 CGCGGGGAGCAAGCAAGCTGCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 Oy 1201 CAACGCAATTAACCGATGAG 1260

Db 1201 CAACGCAATTAACCGATGAG 1260
 Oy 1261 GCGCCGAGCCCTGTCGCGGTCATCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Db 1261 GCGCCGAGCCCTGTCGCGGTCATCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Oy 1321 TAGCCTCAATGAGACTGCTGATGATCAAGGCTTTGGCAAGACAGAGGAGAGAGAGAGAG 1380
 Db 1321 TAGCCTCAATGAGACTGCTGATGATCAAGGCTTTGGCAAGACAGAGGAGAGAGAGAGAG 1380
 Oy 1381 GACATCCCTCTCTCTCGGAGAGTGCAGAGTCAATCTCATGCACTTCAAGAAATGCAATGA 1440
 Db 1381 GACATCCCTCTCTCTCGGAGAGTGCAGAGTCAATCTCATGCACTTCAAGAAATGCAATGA 1440
 Oy 1441 CTACTGGTCTATGAGAGTTACCTTACCCCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 1441 CTACTGGTCTATGAGAGTTACCTTACCCCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Oy 1501 GGGCAGAGACTCTGCGAG 1560
 Db 1501 GGGCAGAGACTCTGCGAG 1560
 Oy 1561 CTGCTACTGCGAGAGTGCACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG 1620
 Db 1561 CTGCTACTGCGAGAGTGCACAGCTGGGGGACAGGCTGTGGCCAGAGAAACAAACCTGG 1620
 Oy 1621 TGATGACCAACAAAGTGAAGAGAGTCTTCCCTGATTTACAGCAAGATGAGAGAGAGT 1680
 Db 1621 TGATGACCAACAAAGTGAAGAGAGTCTTCCCTGATTTACAGCAAGATGAGAGAGAGT 1680
 Oy 1681 GCGATTGAGAAATCTTAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Db 1681 GCGATTGAGAAATCTTAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 Oy 1741 TCGAGAAA 1748
 Db 1741 TCGAGAAA 1748

RESULT 2
 US-10-806-370-11
 ; Sequence 11, Application US/10806370
 ; Publication No. US20040209327A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xieo, Yonhong
 ; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
 ; FILE REFERENCE: 02973.00035
 ; CURRENT APPLICATION NUMBER: US/10/806,370
 ; CURRENT FILING DATE: 2004-03-23
 ; PRIOR APPLICATION NUMBER: US/09/879,792
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/211,224
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: US 60/283,353
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: US 60/283,648
 ; PRIOR FILING DATE: 2001-04-16
 ; PRIOR APPLICATION NUMBER: PCT _____ (Docket No. LIO-81-WO)
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 1748
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-806-370-11

Query Match 100.0%; Score 1748; DB 9; Length 1748;
 Best Local Similarity 100.0%; Pzed. No. 0;
 Matches 1748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGAGACCATGAGAGAGGAGCAGCGAAGGAAATGATCTTCAGACAGAAACACCTTGAGC 60
DB 1 CTGAGAGACCATGAGAGAGGAGCAGCGAAGGAAATGATCTTCAGACAGAAACACCTTGAGC 60
QY 61 TGGAGCATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 120
DB 61 TGGAGCATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 120
QY 121 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 180
DB 121 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 180
QY 181 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 240
DB 181 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 240
QY 241 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 300
DB 241 CCAGGATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 300
QY 301 CAGGTCATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 360
DB 301 CAGGTCATCTTCAGAGCCAGGAGCATCTCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 360
QY 361 TAGAGCAACCAAGTGGGGGCTGTAACCATCCGATCATCTCCAGCTGGAGACCTTCAGAGC 420
DB 361 TAGAGCAACCAAGTGGGGGCTGTAACCATCCGATCATCTCCAGCTGGAGACCTTCAGAGC 420
QY 421 AACGAGGAGCCAGGAGAGAGCCAGGATCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 480
DB 421 AACGAGGAGCCAGGAGAGAGCCAGGATCCAGCTGGAGACCTTCAGAGCCGGGATCTCCAGC 480
QY 481 CCAAGACAGCTACCGCTCATCGGGTGGTGTCTCTCTCATTTGAGCTGGTGGTCT 540
DB 481 CCAAGACAGCTACCGCTCATCGGGTGGTGTCTCTCTCATTTGAGCTGGTGGTCT 540
QY 541 CATATATCTCTTCAGAGCTTCAGAGGAGCAGACAGGATTCAGAGGAGAGAGAGAGAGAG 600
DB 541 CATATATCTCTTCAGAGCTTCAGAGGAGCAGACAGGATTCAGAGGAGAGAGAGAGAGAG 600
QY 601 GAGCTGTCCCAAGAGAGCTGTGTGCTGAGAGGAGGATTCAGAGCTGGAGAGAGAGAG 660
DB 601 GAGCTGTCCCAAGAGAGCTGTGTGCTGAGAGGAGGATTCAGAGCTGGAGAGAGAGAG 660
QY 661 GAGCTGTCCCAAGAGAGCTGTGTGCTGAGAGGAGGATTCAGAGCTGGAGAGAGAGAG 720
DB 661 GAGCTGTCCCAAGAGAGCTGTGTGCTGAGAGGAGGATTCAGAGCTGGAGAGAGAGAG 720
QY 721 CTCCTCATCATGAGGCTTCCTCATCTGTAGAGAGACCTGAAATGATCTTACTGAGAGAGAG 780
DB 721 CTCCTCATCATGAGGCTTCCTCATCTGTAGAGAGACCTGAAATGATCTTACTGAGAGAGAG 780
QY 781 CTGTCAGAGAGCTGGAGTTCAGAGAGTCTCACCGAGCAACCGAGTGGCCCAAGAGATTT 840
DB 781 CTGTCAGAGAGCTGGAGTTCAGAGAGTCTCACCGAGCAACCGAGTGGCCCAAGAGATTT 840
QY 841 TGCCACAGCTTCTCAATCTTGAGATACATCTCACCATCCAGAGAAAGCTTCCAGAGTTC 900
DB 841 TGCCACAGCTTCTCAATCTTGAGATACATCTCACCATCCAGAGAAAGCTTCCAGAGTTC 900
QY 901 TGAATGCTTCCTCCAGAGGATATCTCCCTCCAGAGTTCCTCCAGAGGATTCAGAGGATTC 960
DB 901 TGAATGCTTCCTCCAGAGGATATCTCCCTCCAGAGTTCCTCCAGAGGATTCAGAGGATTC 960
QY 961 GACCGGAGGATCTGGAGAGGAGGCTGCTCCGATAGCAAGTGGAGCTTGGAGAGTGGAG 1020
DB 961 GACCGGAGGATCTGGAGAGGAGGCTGCTCCGATAGCAAGTGGAGCTTGGAGAGTGGAG 1020
QY 1021 TCTGCACTTGGGACCAACCAATCTGTGAGAGCAGCTCATTTGAGAGGAGGAGTGGAGTCT 1080
DB 1021 TCTGCACTTGGGACCAACCAATCTGTGAGAGCAGCTCATTTGAGAGGAGGAGTGGAGTCT 1080

QY 1081 CACTGCCGCCACTGCTCTTCTGTGACCCGGAGAGAGAGTCTTGAGAGGCTGGAAGTGTGA 1140
DB 1081 CACTGCCGCCACTGCTCTTCTGTGACCCGGAGAGAGAGTCTTGAGAGGCTGGAAGTGTGA 1140
QY 1141 CGGGGAGCAGAGCACTGCAACGATTTGAGAGGAGCTTCCATTGCGAGATCATCAT 1200
DB 1141 CGGGGAGCAGAGCACTGCAACGATTTGAGAGGAGCTTCCATTGCGAGATCATCAT 1200
QY 1201 CAACAGCAATTTACACCGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 CAACAGCAATTTACACCGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 GCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 GCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 TAGCTCAATGAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 TAGCTCAATGAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 GACATCCCCCTCTCTCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 GACATCCCCCTCTCTCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 CTACTGATCTATGAGAGCTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATG 1500
DB 1441 CTACTGATCTATGAGAGCTTACCTTACCCCAAGATGATGATGATGATGATGATGATGATG 1500
QY 1501 GGGGAGAGACTCTGCGAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGAGAGAGAGAGAG 1560
DB 1501 GGGGAGAGACTCTGCGAGGAGAGCAGCGGGGGGCTCTTGTCTGTGAGAGAGAGAGAGAG 1560
QY 1561 CTGCTACCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 CTGCTACCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 TGTGTACACCAAGATGAGAGAGTCTTCCCTGAGATTTACAGCAAGATGAGAGAGAGT 1680
DB 1621 TGTGTACACCAAGATGAGAGAGTCTTCCCTGAGATTTACAGCAAGATGAGAGAGAGT 1680
QY 1681 GCGATTCAGAAATCTTACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 1681 GCGATTCAGAAATCTTACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 TCGAGAAA 1748
DB 1741 TCGAGAAA 1748

RESULT 3
US-10-428-275-129
; Sequence 129, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23

```

1 PRIOR APPLICATION NUMBER: 60/185548
2
3 PRIOR FILING DATE: 2000-02-25
4
5 NUMBER OF SEQ ID NOS: 450
6
7 SOFTWARE: Cribaseqblast version 0.1
8
9 SEQ ID NO: 129
10
11 LENGTH: 1707
12
13 TYPE: DNA
14
15 ORGANISM: Homo sapiens
16
17 FEATURE:
18
19 NAME/KEY: CDS
20
21 LOCATION: (1)..(1707)
22
23 OS=10-428-275-129

```

Query Match	96.6%	Score 1689;	DB 8;	Length 1707;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1689; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	8	ACCATGAGAGGGA	CAGCCACGGGAATGATCTCCAGACAAGAACCTTCAGCTGAGCA	67
Db	10	ACCATGAGAGGGA	CAGCCACGGGAATGATCTCCAGACAAGAACCTTCAGCTGAGCA	69
QY	68	TCTCCAGCCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGGGACATCTCCAGCCGAGCA	127	
Db	70	TCTCCAGCCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGGGACATCTCCAGCCGAGCA	129	
QY	128	TCTCCAGCCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGGGACATCTCCAGCCGAGCA	187	
Db	130	TCTCCAGCCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGGGACATCTCCAGCCGAGCA	189	
QY	188	TCTCCAGCTGTAACAACCTCCAGGCGGGGACATCTCCAGGCGGGGACATCTCCAGCCGAGCA	247	
Db	190	TCTCCAGCTGTAACAACCTCCAGGCGGGGACATCTCCAGGCGGGGACATCTCCAGCCGAGCA	249	
QY	248	TCTCCAGCCCGGAGCATCTCCGGCTCTGGCATCATCTTCAGGTCCTCATCCGAGAGGTCA	307	
Db	250	TCTCCAGCCCGGAGCATCTCCGGCTCTGGCATCATCTTCAGGTCCTCATCCGAGAGGTCA	309	
QY	308	TGATCCGGCAGGTCAGGCTCGGTGGTGAACAACCTCCCAACAAGTGTACCTGTTTAGCA	367	
Db	310	TGATCCGGCAGGTCAGGCTCGGTGGTGAACAACCTCCCAACAAGTGTACCTGTTTAGCA	369	
QY	368	ACACAGTGGGGCTGTACCACTCCGATCATCTCTCCAGGTGAGCAACCAAGCAACGAG	427	
Db	370	ACACAGTGGGGCTGTACCACTCCGATCATCTCTCCAGGTGAGCAACCAAGCAACGAG	429	
QY	428	GCCACCAAGGAGAGCCAGAGTACAGAGCTGCCAAGTTACCTCGCGGAGAGGACAGAG	487	
Db	430	GCCACCAAGGAGAGCCAGAGTACAGAGCTGCCAAGTTACCTCGCGGAGAGGACAGAG	489	
QY	488	CAGCTACCGGCTCATCGGGTGGGTGCTCCCTCCATATGGCCGTGGTGGTTTGCTCATATC	547	
Db	490	CAGCTACCGGCTCATCGGGTGGGTGCTCCCTCCATATGGCCGTGGTGGTTTGCTCATATC	549	
QY	548	CTCTTCAGTTCGTGAGAGGCCACAAGGATCAGGTACAAAGAGCAGAAGGAGAGCTGT	607	
Db	550	CTCTTCAGTTCGTGAGAGGCCACAAGGATCAGGTACAAAGAGCAGAAGGAGAGCTGT	609	
QY	608	CCCAACGACGCGTTCGTGCTGTGAACGGGGTGGTGTGACCTGCAAGCTGGAAGTGAAGAGCTG	667	
Db	610	CCCAACGACGCGTTCGTGCTGTGAACGGGGTGGTGTGACCTGCAAGCTGGAAGTGAAGAGCTG	669	
QY	668	GGCTGGGTGAGGTTTGTACTGGGACAAAGTCTGCTTAATAATCTACTCGGGTCTCTCCAT	727	
Db	670	GGCTGGGTGAGGTTTGTACTGGGACAAAGTCTGCTTAATAATCTACTCGGGTCTCTCCAT	729	
QY	728	CAGTGGCTTCCATCTGTAGACAGCACTGGATGATCTCTTACTGAGAGAGACTTGCAG	787	
Db	730	CAGTGGCTTCCATCTGTAGACAGCACTGGATGATCTCTTACTGAGAGAGACTTGCAG	789	
QY	788	CAGCTGGGTTTGAAGAGTCTCACCGGACAAACCGAGTTGCCACAGGAGATTTTGCCAAC	847	
Db	790	CAGCTGGGTTTGAAGAGTCTCACCGGACAAACCGAGTTGCCACAGGAGATTTTGCCAAC	849	

Qy	848	AGCTTCATCTTTGAGATACATCAACCCACATCCAGGAAGAGCTCCACAGGTCGAATTGC	907
Dp	850	AGCTTCATCTTTGAGATACATCTTCAACATCCAGGAAGAGCTCCACAGGTCGAATTGC	909
Qy	908	CCTTCCAGCGGATATATCTCCCTCAGTGTTCACCTGCGGACTGAGGCCATGACCCGG	967
Dp	910	CCTTCCAGCGGATATATCTCCCTCAGTGTTCACCTGCGGACTGAGGCCATGACCCGG	969
Qy	968	CGGATCGTGGGAGGGGGGCGTGGCCCTGGGATATGCAAGTGGCTTGGCAAGTGAAGTCTCAC	1027
Dp	970	CGGATCGTGGGAGGGGGGCGTGGCCCTGGGATATGCAAGTGGCTTGGCAAGTGAAGTCTCAC	1029
Qy	1028	TTCCGGCAACAACCCATCTGTGAGGAGCGCTCATTTGACGCCACGAGTGGTCTCATCTGCC	1087
Dp	1030	TTCCGGCAACAACCCATCTGTGAGGAGCGCTCATTTGACGCCACGAGTGGTCTCATCTGCC	1089
Qy	1088	GCCCATCTGCTTTCTGTGACCCCGGAGAAAGTCTCTGAGGGCTTGGAAAGTGTACGCGGC	1147
Dp	1090	GCCCATCTGCTTTCTGTGACCCCGGAGAAAGTCTCTGAGGGCTTGGAAAGTGTACGCGGC	1149
Qy	1148	ACGAGCAACCTGCACCAAGTTGCTCTGAGGACAGCCTCCATTGCCGAGATCATCATCAAGC	1207
Dp	1150	ACGAGCAACCTGCACCAAGTTGCTCTGAGGACAGCCTCCATTGCCGAGATCATCATCAAGC	1209
Qy	1208	AATTACACCGATGAGAGAGACGACTATGACATGCGCCTCATGCGGCTGTCCAAAGCCCTG	1267
Dp	1210	AATTACACCGATGAGAGAGACGACTATGACATGCGCCTCCATGCGGCTGTCCAAAGCCCTG	1269
Qy	1268	ACCGTGCCTGCTCAATCCACCCGCTTGGCTCCGCCATGAGATGGACAGACCTTATAGCCTC	1327
Dp	1270	ACCGTGCCTGCTCAATCCACCCGCTTGGCTCCGCCATGAGATGGACAGACCTTATAGCCTC	1329
Qy	1328	AATGAGACCTGCTGATACAGAGCTTTGGCAAGACCGAGAGACAGATGACAAAGACATCC	1387
Dp	1330	AATGAGACCTGCTGATACAGAGCTTTGGCAAGACCGAGAGACAGATGACAAAGACATCC	1389
Qy	1388	CCCTTCTCCGGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG	1447
Dp	1390	CCCTTCTCCGGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG	1449
Qy	1448	GTCATGACAGTTACCTTACCCCAAGGATGATGTCCTGGGACCTTCTGCGGGGAGAGA	1507
Dp	1450	GTCATGACAGTTACCTTACCCCAAGGATGATGTCCTGGGACCTTCTGCGGGGAGAGA	1509
Qy	1508	GACTCTGCGACAGGAGACAGCCGGGGGCTCTTGTCTGTGAGCAAGAACCAACCTGTGTAC	1567
Dp	1510	GACTCTGCGACAGGAGACAGCCGGGGGCTCTTGTCTGTGAGCAAGAACCAACCTGTGTAC	1569
Qy	1568	CTGGCAGGTTCACCAACTGTGGGACAGGCGCTGTGGCCAGAGAAACAAACCTGGTGTAC	1627
Dp	1570	CTGGCAGGTTCACCAACTGTGGGACAGGCGCTGTGGCCAGAGAAACAAACCTGGTGTAC	1629
Qy	1628	ACCAAAATGACAGAAAGTTCTTCCCTGTGATTTTACAGCAAGATGGAGACGAGGTGCAATTC	1687
Dp	1630	ACCAAAATGACAGAAAGTTCTTCCCTGTGATTTTACAGCAAGATGGAGACGAGGTGCAATTC	1689
Qy	1688	AGAAATATCC	1696
Dp	1690	AGAAATATCC	1698

```

RESULT 4
US-10-428-275-147/c
; Sequence 147, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
;

```


RESULT 5
US-10-156-214A-23
; Sequence 23, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempke
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siew
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Theofel
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)...(1689)
; OTHER INFORMATION: Nucleic acid encoding Endothelialase 2-S protein
US-10-156-214A-23

Query Match 96.5%; Score 1687.4; DB 7; Length 1689;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 ATGGAGAGGACAGCCACGGGATGCACTTCCAGCAAGAACCTTCAGCTGAGCATCT 70
Db 1 ATGGAGAGGACAGCCACGGGATGCACTTCCAGCAAGAACCTTCAGCTGAGCATCT 60
QY 71 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 130
Db 61 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 120
QY 131 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 190
Db 121 CCAGCCCAAGGATCTCCAGCTGGGACACCTTCAGAGCCGGGATCTCCAGCCAGCATCT 180
QY 191 CCAGCTGTACACCTTCAGAGCCGGGACATCTTCAGAGCCGGGATCTCCAGCCAGCATCT 250
Db 181 CCAGCTGTACACCTTCAGAGCCGGGACATCTTCAGAGCCGGGATCTCCAGCCAGCATCT 240
QY 251 CCAGCCCGGGATCTCCAGCTGGGACATCTTCAGAGCTTCATTCAGAGCTTCATCA 310
Db 241 CCAGCCCGGGATCTCCAGCTGGGACATCTTCAGAGCTTCATTCAGAGCTTCATCA 300
QY 311 TCCGCGAGGTCAGCTCGGTGACAACTCCCAACAGAGTGAACCTTGTAGAGCAACA 370
Db 301 TCCGCGAGGTCAGCTCGGTGACAACTCCCAACAGAGTGAACCTTGTAGAGCAACA 360
QY 371 CCAAGTGGAGGCTGTACCATTCGATCACTTCCTGCAAGTTCAGAGCAAGCAAGGAGCC 430
Db 361 CCAAGTGGAGGCTGTACCATTCGATCACTTCCTGCAAGTTCAGAGCAAGCAAGGAGCC 420
QY 431 ACCAGGGAGAGCCCAAGTGAAGAGCTGCGCAAGTTCACTTGGGAGAGGAGCAAGAGAG 490
Db 421 ACCAGGGAGAGCCCAAGTGAAGAGCTGCGCAAGTTCACTTGGGAGAGGAGCAAGAGAG 480
QY 491 CTACCGCTCATCGGAGTGGTCTCTCTCATTTGCGCTGAGTGTGGCTCATCTCTC 550
Db 481 CTACCGCTCATCGGAGTGGTCTCTCTCATTTGCGCTGAGTGTGGCTCATCTCTC 540
QY 551 TTCCAGTTCTTGAGAGGAGCAACAGGATCAAGTGAACAAGAGAGAGAGAGTGTCTCC 610
Db 541 TTCCAGTTCTTGAGAGGAGCAACAGGATCAAGTGAACAAGAGAGAGAGTGTCTCC 600
QY 611 AAGCAGCCTGTTCGCTGTGACGGGAGTGGAGCTGCAAGCTGAAGAGTGAAGCTGGCC 670

Db 601 AAGCAGCCTGTTCGCTGTGACGGGAGTGGAGCTGCAAGCTGAAGAGTGAAGCTGGCC 660
QY 671 TGCCTGAGGTTTGAAGTGGAGCAAGTCTCTGCTTAAATCTACTCTGGAGCTCCCATACAG 730
Db 661 TGCCTGAGGTTTGAAGTGGAGCAAGTCTCTGCTTAAATCTACTCTGGAGCTCCCATACAG 720
QY 731 TGCCTTCCCATCTGTAGAGCAAGTGAAGTCTCTACTGAGAGAAAGTCTGCAAGAG 790
Db 721 TGCCTTCCCATCTGTAGAGCAAGTGAAGTCTCTACTGAGAGAAAGTCTGCAAGAG 780
QY 791 CTGGGTTTCCAGAGTGTCAACCGGACCAAGAGTTTCCCAAGAGATTTTCCCAAGAGC 850
Db 781 CTGGGTTTCCAGAGTGTCAACCGGACCAAGAGTTTCCCAAGAGATTTTCCCAAGAGC 840
QY 851 TTCTCAATCTTGAAGATCAACTCCAGCATCAGGAAAGCTCCAGAGTGAATGAGCTCT 910
Db 841 TTCTCAATCTTGAAGATCAACTCCAGCATCAGGAAAGCTCCAGAGTGAATGAGCTCT 900
QY 911 TCCAGCGGATATATCTCCCTCCAGTGTTCCTCACTGGCGACTGAGAGGCTATGACCGGCGG 970
Db 901 TCCAGCGGATATATCTCCCTCCAGTGTTCCTCACTGGCGACTGAGAGGCTATGACCGGCGG 960
QY 971 ATCTGGAGAGGAGGCTGGCTGGAGTGAAGAGGCTTGGCAAGTGAAGTCTGCACTTC 1030
Db 961 ATCTGGAGAGGAGGCTGGCTGGAGTGAAGAGGCTTGGCAAGTGAAGTCTGCACTTC 1020
QY 1031 GGCACCAACCAATCTGTGAGAGGACGCTCATTTGACCGCCAGTGGAGTGTCACTGCGGCC 1090
Db 1021 GGCACCAACCAATCTGTGAGAGGACGCTCATTTGACCGCCAGTGGAGTGTCACTGCGGCC 1080
QY 1091 CACTGCTTCTTGTGACCCGCGAGAGAGGCTCTGAGAGGCTTGGAAAGTGTACCGGCGACCC 1150
Db 1081 CACTGCTTCTTGTGACCCGCGAGAGAGGCTCTGAGAGGCTTGGAAAGTGTACCGGCGACCC 1140
QY 1151 AGCAACCTTGACAGGTTGCTTGAAGGAGCTTCCATTCGCGAATCATCATCAAGCAAT 1210
Db 1141 AGCAACCTTGACAGGTTGCTTGAAGGAGCTTCCATTCGCGAATCATCATCAAGCAAT 1200
QY 1211 TACACGATGAGAGAGAGCACTATGATGATGAGCCCTTCATGCGGCTGTCAAGCCCTGACC 1270
Db 1201 TACACGATGAGAGAGAGCACTATGATGATGAGCCCTTCATGCGGCTGTCAAGCCCTGACC 1260
QY 1271 CTGTGCTCATCATCAACCTCTGTGCTTCCCATGATGAGACAGACCTTTAGCTCAAT 1330
Db 1261 CTGTGCTCATCATCAACCTCTGTGCTTCCCATGATGAGACAGACCTTTAGCTCAAT 1320
QY 1331 GAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGAGCAAGTGAACAAGACATCCGCC 1390
Db 1321 GAGACCTGTGATCAAGGCTTTGGCAAGCAAGGAGAGCAAGTGAACAAGACATCCGCC 1380
QY 1391 TTCTCCGAGAGGTCAGGTCATCTCATGACATTCGAAGAAATGCAATGACTTGGTTC 1450
Db 1381 TTCTCCGAGAGGTCAGGTCATCTCATGACATTCGAAGAAATGCAATGACTTGGTTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCGGAGACTTTCGTGGGAGCAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGATGATGTGTGCGGAGACTTTCGTGGGAGCAGAGAC 1500
QY 1511 TCCGCGAGGAGACAGCGGGGGGCTCTGTGTGTGAGCAAGCAACCGCTGGTACCTG 1570
Db 1501 TCCGCGAGGAGACAGCGGGGGGCTCTGTGTGTGAGCAAGCAACCGCTGGTACCTG 1560
QY 1571 GCAAGTGTCAACAGCTGGGAGCAGGCTGTGTGCGCAGAGAAACAACTGTGTGTACAC 1630
Db 1561 GCAAGTGTCAACAGCTGGGAGCAGGCTGTGTGCGCAGAGAAACAACTGTGTGTACAC 1620
QY 1631 AAGTGAACAAGTTCTTCCCTGATTTTACAGCAAGTGAAGAGAGGAGTGCATTGAGA 1690
Db 1621 AAGTGAACAAGTTCTTCCCTGATTTTACAGCAAGTGAAGAGAGGAGTGCATTGATA 1680
QY 1691 AATCTCTAA 1699
|||||

Db 1681 AAATCTTA 1689

RESULT 6
US-10-916-758-93
; Sequence 93, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-916-758-93

Query Match 96.5%; Score 1687.4; DB 10; Length 1689;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 ATGGAGAGGACACGACGAGGATGATCTCTCAGCAAGAACACCTTCAGTGAAGCATCT 70
Db 1 ATGGAGAGGACACGACGAGGATGATCTCTCAGCAAGAACACCTTCAGTGAAGCATCT 60
QY 71 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 130
Db 61 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 120
QY 131 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 190
Db 121 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 180
QY 191 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 250
Db 181 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 240
QY 251 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 310
Db 241 CCAAGCCAGGATCTCAGCTGGAGCACTTCAGAGCCGGGATCTCCAGCCAGGATCT 300
QY 311 TCCGCGAGTACAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTTAGACAA 370
Db 301 TCCGCGAGTACAGCTCGGTGACAACTCCCAACAGAGTACCTTTGTTAGACAA 360
QY 371 CCAAGTGGGGGCTGATCCATCCATCTCTGCAAGTCTTCAGAGTCTTCAGAGTCT 430
Db 361 CCAAGTGGGGGCTGATCCATCCATCTCTGCAAGTCTTCAGAGTCTTCAGAGTCT 420
QY 431 ACCAGGAGAGCCAGGATCGAGCTGCGCAAGTTCACTTGGCGGAGAGGCGCAAGAG 490
Db 421 ACCAGGAGAGCCAGGATCGAGCTGCGCAAGTTCACTTGGCGGAGAGGCGCAAGAG 480
QY 491 CTACCGCTCATCGGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 550
Db 481 CTACCGCTCATCGGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 551 TTCCAGTCTTGGAGGAGGACACAGAGGATCAGGTACAAAGAGAGAGAGGAGGAGCTGCC 610
Db 541 TTCCAGTCTTGGAGGAGGACACAGAGGATCAGGTACAAAGAGAGAGGAGGAGCTGCC 600
QY 611 AAGCAGCGCTGCTCTGTGACGGGGTGGTGGAGCTGCAAGCTGAAGAGTGAAGAGCTGGC 670
Db 601 AAGCAGCGCTGCTCTGTGACGGGGTGGTGGAGCTGCAAGAGTGAAGAGTGAAGAGCTGGC 660

QY 671 TGGTGAGGTTTGAAGTGAAGTCTCTGTTAAATCTACTGAGGCTCTCCATCAG 730
Db 661 TGGTGAGGTTTGAAGTGAAGTCTCTGTTAAATCTACTGAGGCTCTCCATCAG 720
QY 731 TGGTCTCCATCTGTAGACCAACTGGAATGACTCTTACTGAGAGAGACTGCGACAG 790
Db 721 TGGTCTCCATCTGTAGACCAACTGGAATGACTCTTACTGAGAGAGACTGCGACAG 780
QY 791 CTGGGTTTCAAGAGTGTACCCGAGCAACCGAGGTTTCCACAGAGGATTTTCCAAACAG 850
Db 781 CTGGGTTTCAAGAGTGTACCCGAGCAACCGAGGTTTCCACAGAGGATTTTCCAAACAG 840
QY 851 TTCTCAATCTTGAATCACTCCACCATCAGAGAAAGCTCCACAGGTCGAATGCGCT 910
Db 841 TTCTCAATCTTGAATCACTCCACCATCAGAGAAAGCTCCACAGGTCGAATGCGCT 900
QY 911 TCCAGCGGATATCTCCCTCCAGTGTTCCTCAGTGGAGTGAAGGATGATGACCGGCGG 970
Db 901 TCCAGCGGATATCTCCCTCCAGTGTTCCTCAGTGGAGTGAAGGATGATGACCGGCGG 960
QY 971 ATGTGGAGAGGAGGCTGAGTGGATGAGCAAGTGGCTTGGCAAGTGAAGTTCATTC 1030
Db 961 ATGTGGAGAGGAGGCTGAGTGGATGAGCAAGTGGCTTGGCAAGTGAAGTTCATTC 1020
QY 1031 GGCACCAACCCATCTGTGAGAGGACGCTCATTTGACGCCAGTGGAGTCACTGCGGCC 1090
Db 1021 GGCACCAACCCATCTGTGAGAGGACGCTCATTTGACGCCAGTGGAGTCACTGCGGCC 1080
QY 1091 CACTGCTTCTTGTGACCCCGGAGAAAGTCTGTGAGAGGCTGGAAGGTGTACCGCGGACC 1150
Db 1081 CACTGCTTCTTGTGACCCCGGAGAAAGTCTGTGAGAGGCTGGAAGGTGTACCGCGGACC 1140
QY 1151 AGCAACTGCAACAGTTCCTGAGGAGGCTCTCCATTCGCGAGATCATATCAACAGCAAT 1210
Db 1141 AGCAACTGCAACAGTTCCTGAGGAGGCTCTCCATTCGCGAGATCATATCAACAGCAAT 1200
QY 1211 TACACGATGAGAGGAGCACTATGATGAGCCCTTACGCGGCTGTCAAGCCCTGAGCC 1270
Db 1201 TACACGATGAGAGGAGCACTATGATGAGCCCTTACGCGGCTGTCAAGCCCTGAGCC 1260
QY 1271 CTGTCCGCTCAATCCACCCCTGCTTGGCTCCCATGAGATGAGACGACCTTATAGCTCAAT 1330
Db 1261 CTGTCCGCTCAATCCACCCCTGCTTGGCTCCCATGAGATGAGACGACCTTATAGCTCAAT 1320
QY 1331 GAGACTGCTGATCAGAGCTTTGGCAAGACAGGAGACAGATGACAGACATCCCCC 1390
Db 1321 GAGACTGCTGATCAGAGCTTTGGCAAGACAGGAGACAGATGACAGACATCCCCC 1380
QY 1391 TTCTCCGAGAGGAGGATCAATCTGATGACCTTCAAGAAATGCAATGACTTATGCTGTC 1450
Db 1381 TTCTCCGAGAGGAGGATCAATCTGATGACCTTCAAGAAATGCAATGACTTATGCTGTC 1440
QY 1451 TATGACAGTTACCTTACCCCAAGAGATGATGTCCTGGGACCTTCTGAGGAGGAGAGAC 1510
Db 1441 TATGACAGTTACCTTACCCCAAGAGATGATGTCCTGGGACCTTCTGAGGAGGAGAGAC 1500
QY 1511 TCTGCGAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGAGAGAAACACCGTGTGACTG 1570
Db 1501 TCTGCGAGGAGAGACAGCGGGGGCTCTTGTCTGTGAGAGAGAAACACCGTGTGACTG 1560
QY 1571 GCGAGTGTCAACAGCTGGGAGACAGGCTGTGGCCAGAGAGAAACACCGTGTGTACACC 1630
Db 1561 GCGAGTGTCAACAGCTGGGAGACAGGCTGTGGCCAGAGAGAAACACCGTGTGTACACC 1620
QY 1631 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGGAGTGCATTCAGA 1690
Db 1621 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGAGGAGTGCATTCATA 1680
QY 1691 AAATCTTA 1699
Db 1681 AAATCTTA 1689

RESULT 7

US-10-353-690-99

/ Sequence 99, Application US/10353690
/ Publication No. US20030215840A1

/ GENERAL INFORMATION:

/ APPLICANT: Logan, Thomas Joseph

/ APPLICANT: Chun, Miyoung

/ APPLICANT: Galvin, Katherine M.

/ APPLICANT: Healy, Aileen

/ APPLICANT: Accorn, Susan L.

/ APPLICANT: Donoghue, Mary

/ APPLICANT: Stegiano, Nancy

/ APPLICANT: Perodine, Jacqueline

/ APPLICANT: Rodriguez-Way, Amelie

/ TITLE OF INVENTION: Methods and compositions for treating

/ TITLE OF INVENTION: Cardiovascular disease using 1682, 6169, 6193, 7771, 14395,

/ TITLE OF INVENTION: 29002, 3216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,

/ TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,

/ TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,

/ TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,

/ TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,

/ TITLE OF INVENTION: 51130, 19469, 21833, 2917, 59590, 15992, 2094, 2252, 3474,

/ FILE REFERENCE: MP102-018P1RONMINI

/ CURRENT APPLICATION NUMBER: US/10/353,690

/ PRIOR APPLICATION NUMBER: 60/353,224

/ PRIOR FILING DATE: 2002-02-01

/ PRIOR APPLICATION NUMBER: 60/364,529

/ PRIOR FILING DATE: 2002-03-15

/ PRIOR APPLICATION NUMBER: 60/373,861

/ PRIOR FILING DATE: 2002-04-19

/ PRIOR APPLICATION NUMBER: 60/376,287

/ PRIOR FILING DATE: 2002-04-29

/ PRIOR APPLICATION NUMBER: 60/388,080

/ PRIOR FILING DATE: 2002-06-12

/ PRIOR APPLICATION NUMBER: 60/390,971

/ PRIOR FILING DATE: 2002-06-24

/ PRIOR APPLICATION NUMBER: 60/394,130

/ PRIOR FILING DATE: 2002-07-03

/ PRIOR APPLICATION NUMBER: 60/394,797

/ PRIOR FILING DATE: 2002-07-10

/ PRIOR APPLICATION NUMBER: 60/404,904

/ PRIOR FILING DATE: 2002-08-21

/ PRIOR APPLICATION NUMBER: 60/405,450

/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 126

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 99

/ LENGTH: 2393

/ TYPE: DNA

/ ORGANISM: Homo Sapiens

/ US-10-353-690-99

Query Match 95.6%; Score 1670.4; DB 7; Length 2393;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCAGAGACCATGAGAGGAGACAGCCAGGAGATGATCTTCACAGAGAACCTTCAGC 60
DB 78 CTCAGAGACCATGAGAGGAGAGACAGCCAGGAGATGATCTTCACAGAGAACCTTCAGC 137
QY 61 TGGAGCATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 120
DB 138 TGGAGCATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 197
QY 121 CCAAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 180
DB 198 CCAAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 257
QY 181 CCAAGGATCTCCAGCTGGATACCTTCAGCCGGGAGATCTTCAGCCGGGAGATCTTCAGC 240

DB 258 CCAAGGATCTCCAGCTGGATACCTTCAGCCGGGAGATCTTCAGCCGGGAGATCTTCAGC 317
QY 241 CCAAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 300
DB 318 CCAAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 377
QY 301 CAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 360
DB 378 CAGGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 437
QY 361 TAGAGCAACCAAGTGGGGGCTGATCCATCCATCTTCAGCTGGAGACACTTCAGC 420
DB 438 TAGAGCAACCAAGTGGGGGCTGATCCATCCATCTTCAGCTGGAGACACTTCAGC 497
QY 421 AACCAAGGAGCCAGGAGAGAGCCAGGATGAGAGCTTCAGCCAGGAGATCTTCAGC 480
DB 498 AACCAAGGAGCCAGGAGAGAGCCAGGATGAGAGCTTCAGCCAGGAGATCTTCAGC 557
QY 481 CCAAGGAGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 540
DB 558 CCAAGGAGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 617
QY 541 CATATCTCTTCAGATTTCTGGAGGAGCCACAGGATGAGATGAGTACAGAGAGAGAGAG 600
DB 618 CATATCTCTTCAGATTTCTGGAGGAGCCACAGGATGAGATGAGTACAGAGAGAGAGAG 677
QY 601 GAGGTGCCAG 660
DB 678 GAGGTGCCAG 737
QY 661 CCAAGGAGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 720
DB 738 CCAAGGAGATCTCCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 797
QY 721 CTCCATCTGAGTGGCTTCATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 798 CTCCATCTGAGTGGCTTCATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
QY 781 CTCCATCTGAGTGGCTTCATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 858 CTCCATCTGAGTGGCTTCATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 917
QY 841 TGCCAGAGATCTTCAGATCTTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 918 TGCCAGAGATCTTCAGATCTTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 977
QY 901 TGAATGCCCTTCCAGAGGATATCTTCCTTCAGATGAGAGAGAGAGAGAGAGAGAG 960
DB 978 TGAATGCCCTTCCAGAGGATATCTTCCTTCAGATGAGAGAGAGAGAGAGAGAGAG 1037
QY 961 GAGCGGGGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1038 GAGCGGGGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
QY 1021 TCTGACATCTGGAG 1080
DB 1098 TCTGACATCTGGAG 1157
QY 1081 CACTGCCGCCCTGCTTCTTCGTGACCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1158 CACTGCCGCCCTGCTTCTTCGTGACCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1217
QY 1141 CGCGGGAG 1200
DB 1218 CGCGGGAG 1277
QY 1201 CCAAGGATCTTCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 1260
DB 1278 CCAAGGATCTTCAGCCAGGAGATCTTCAGCTGGAGACACTTCAGCCGGGAGATCTTCAGC 1337
QY 1261 GCCCTGAGCCCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

Db 1338 GCCCTGACCCCTGTCGCGTCATCATCAACCTTGCTTGCCCTCCCAATGATGACAGACTT 1397
Qy 1321 TAGCTTCAATGAGACCTGCTGATCATACAGGCTTTGGCAAGACAGGGAGACAGATGACAA 1380
Db 1398 TAGCTTCAATGAGACCTGCTGATCATACAGGCTTTGGCAAGACAGGGAGACAGATGACAA 1457
Qy 1381 GACATCCCTCCCTCCGCGGAGGTGACAGGTCAATCTCATGCACTTCAAGAAATGCAATGA 1440
Db 1458 GACATCCCTCCCTCCGCGGAGGTGACAGGTCAATCTCATGCACTTCAAGAAATGCAATGA 1517
Qy 1441 CTACTTGGTCTATGACAGTTACTTATCCCAAGATGATGTGCTGCGGACCTTCGTGG 1500
Db 1518 CTACTTGGTCTATGACAGTTACTTATCCCAAGATGATGTGCTGCGGACCTTCGTGG 1577
Qy 1501 GGGCAGAGACTCTGCGCAGGAGACAGCGGGGGCTTTGTCTGTGAGCAAAACACCG 1560
Db 1578 GGGCAGAGACTCTGCGCAGGAGACAGCGGGGGCTTTGTCTGTGAGCAAAACACCG 1637
Qy 1561 CTGCTACCTGCGCAGGTGTCACAGCTGCGGACAGGCGTGGCCAGAGAAACAAACCTGG 1620
Db 1638 CTGCTACCTGCGCAGGTGTCACAGCTGCGGACAGGCGTGGCCAGAGAAACAAACCTGG 1697
Qy 1621 TGTGTACACCAAGTGAAGAGTTCTTCCCTGATTTTACAGCAAGATGAG 1672
Db 1698 TGTGTACACCAAGTGAAGAGTTCTTCCCTGATTTTACAGCAAGATGAG 1749

RESULT 8
US-10-156-214A-25
; Sequence 25, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempile
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Varma Siev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2067)
; OTHER INFORMATION: Nucleic acid encoding (endothelinase 2-L) protein
US-10-156-214A-25

Query Match 95.2%; Score 1663.4; DB 7; Length 2067;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 ATGAGAGGGGAGCAGCAGGGAATGCATCTCCAGCAAGAACCTTGCAGTGGAGCATCT 70
Db 1 ATGAGAGGGGAGCAGCAGGGAATGCATCTCCAGCAAGAACCTTGCAGTGGAGCATCT 60
Qy 71 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 130
Db 61 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 120
Qy 131 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 190
Db 121 CCAGCCCAAGCATCTCCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 180
Qy 191 CCAGCTGTACACCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 250
Db 1261 CCAGCTGTACACCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 1320

Db 181 CCAAGTGTACACCTCCAGGCCCGGGCATCTCCAGGCCCGGGCATCTCCAGCCCAAGCATCT 240
Qy 251 CCAGCCCGGGCATCTCCGGCTCTGGGATCACTTTCAGAGTCTCATATCCGAGGATCATCA 310
Db 241 CCAGCCCGGGCATCTCCGGCTCTGGGATCACTTTCAGAGTCTCATATCCGAGGATCATCA 300
Qy 311 TCCGCAAGTCTCAGCTCGGTGACAACTCCCAACAGAGTGTACTTTTATGAGCAACA 370
Db 301 TCCGCAAGTCTCAGCTCGGTGACAACTCCCAACAGAGTGTACTTTTATGAGCAACA 360
Qy 371 CCAAGTGGGGCTGTACCTATCCATCCATCTCTCCAGGTGACGACCAAGAGGACCTG 430
Db 361 CCAAGTGGGGCTGTACCTATCCATCCATCTCTCCAGGTGACGACCAAGAGGACCTG 420
Qy 431 ACCAGGAGAGCCCAAGTACGAGCTTCCCAAGTTTACCTGGCGGAGGGCCAAAGAG 490
Db 421 ACCAGGAGAGCCCAAGTACGAGCTTCCCAAGTTTACCTGGCGGAGGGCCAAAGAG 480
Qy 491 CTACCGCTCATCGGGTGGTGGTCTCTCTCTCATTTGCTGGTGTGTTGCTCATCATCTC 550
Db 481 CTACCGCTCATCGGGTGGTGGTCTCTCTCTCATTTGCTGGTGTGTTGCTCATCATCTC 540
Qy 551 TTCCAGTTCTGGCAGGGCCACAGGGATCAGGTACAAAGAGCAGAGGAGAGCTGTCC 610
Db 541 TTCCAGTTCTGGCAGGGCCACAGGGATCAGGTACAAAGAGCAGAGGAGAGCTGTCC 600
Qy 611 AAGCAGCTGTCTGCTGTGACGGGGTGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGCC 670
Db 601 AAGCAGCTGTCTGCTGTGACGGGGTGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGCC 660
Qy 671 TGGGTGAGTTTGAAGTGGGCAAGTCTCTGCTTAAATCTACTCTGGGCTCTCCATCAG 730
Db 661 TGGGTGAGTTTGAAGTGGGCAAGTCTCTGCTTAAATCTACTCTGGGCTCTCCATCAG 720
Qy 731 TGGCTTCCATCTGTAGCAGCAAGTGAATGACTCTTACAGAGAAAGCTGCTCAGAG 790
Db 721 TGGCTTCCATCTGTAGCAGCAAGTGAATGACTCTTACAGAGAAAGCTGCTCAGAG 780
Qy 791 CTGGGTTTCCAGAGTGTCTACCCGACAAACGAGGTGGCCACAGGATTTTGGCCAAACGC 850
Db 781 CTGGGTTTCCAGAGTGTCTACCCGACAAACGAGGTGGCCACAGGATTTTGGCCAAACGC 840
Qy 851 TTCTCAATCTTGAATCAACTCCACCATCCAGGAAAGCTCCAGAGTCTGAATGCCCT 910
Db 841 TTCTCAATCTTGAATCAACTCCACCATCCAGGAAAGCTCCAGAGTCTGAATGCCCT 900
Qy 911 TCCAGCGGTATATCTCCCTCCAGTGTTCACATGCGGATGAGGGCCATGACCGGGGG 970
Db 901 TCCAGCGGTATATCTCCCTCCAGTGTTCACATGCGGATGAGGGCCATGACCGGGGG 960
Qy 971 ATCGTGGAGGGGCTGCGCTCGGATGACAGTGGCTTGGCAAGTGTGCACTTC 1030
Db 961 ATCGTGGAGGGGCTGCGCTCGGATGACAGTGGCTTGGCAAGTGTGCACTTC 1020
Qy 1031 GGCACCAACCAATCTGTGAGGAGCGCTCATTTAGCGCCAGTGGGGTCTCACTGCGGCC 1090
Db 1021 GGCACCAACCAATCTGTGAGGAGCGCTCATTTAGCGCCAGTGGGGTCTCACTGCGGCC 1080
Qy 1091 CACTGCTTCTTCTGTACCCCGGAGAGGTCTCTGAGGGCTGGAAGGTGTACGCGGGCAC 1150
Db 1081 CACTGCTTCTTCTGTACCCCGGAGAGGTCTCTGAGGGCTGGAAGGTGTACGCGGGCAC 1140
Qy 1151 AGCAACTGTGACAGTTGCTGAGGAGCGCTTCATTGCGAGATCATCAACAGCAAT 1210
Db 1141 AGCAACTGTGACAGTTGCTGAGGAGCGCTTCATTGCGAGATCATCAACAGCAAT 1200
Qy 1211 TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAACCTTGAC 1270
Db 1201 TACACCGATGAGAGAGCACTATGACATGCGCCCTCATGCGGCTGTCCAAACCTTGAC 1260
Qy 1271 CTGTCCGCTCATCAACCTGCTGCTCCATGCAATGACAGACCTTATAGCTCAAT 1330
Db 1261 CTGTCCGCTCATCAACCTGCTGCTCCATGCAATGACAGACCTTATAGCTCAAT 1320

Db 1302 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGGAGACGATATGACATCGCCCT 1361
Qy 1246 CATCGGCTGTCCAAAGCCCTGACCCCTGTCCGCTCACATCCACCCCTGCTCCCAT 1305
Db 1362 CATCGGCTGTCCAAAGCCCTGACCCCTGTCCGCTCACATCCACCCCTGCTCCCAT 1421
Qy 1306 GCATGACAGACCTTTAGCCTTCAATGAGACCTGCTGATCA CAGGCTTTGGCAAGACCG 1365
Db 1422 GCATGACAGACCTTTAGCCTTCAATGAGACCTGCTGATCA CAGGCTTTGGCAAGACCG 1481
Qy 1366 GGAGACAGATGACAAAGACATCCCTTCCTCCGGAAGGTGAGATGCAATTCATGACTT 1425
Db 1482 GGAGACAGATGACAAAGACATCCCTTCCTCCGGAAGGTGAGATGCAATTCATGACTT 1541
Qy 1426 CAAGAATGCAATGACTTCTGTATGACAGTTACTTCCCAAGATGATGTGC 1485
Db 1542 CAAGAATGCAATGACTTCTGTATGACAGTTACTTCCCAAGATGATGTGC 1601
Qy 1486 TGGGACCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGGCTTTGTCTG 1545
Db 1602 TGGGACCTTGTGGGGGAGAGACTCTGCGAGGAGACAGCGGGGGGCTTTGTCTG 1661
Qy 1546 TGACGAGAAACAACCGCTGTGACCTGCGAGGTGTACACGAGCTGGGGGACAGGCTGTGCGCA 1605
Db 1662 TGACGAGAAACAACCGCTGTGACCTGCGAGGTGTACACGAGCTGGGGGACAGGCTGTGCGCA 1721
Qy 1606 GAGAAACAACCTGTGTGTACACCAAGTGTGACAGAAAGTTCTTCCCTGATTTACAGCA 1665
Db 1722 GAGAAACAACCTGTGTGTGTACACCAAGTGTGACAGAAAGTTCTTCCCTGATTTACAGCA 1781
Qy 1666 GATGAG 1672
Db 1782 GATGAG 1788

RESULT 11

US-10-428-275-139
Sequence 139, Application US/10428275
Publication No. US20040067505A1

GENERAL INFORMATION:
APPLICANT: Alvarez et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585

CURRENT APPLICATION NUMBER: US/10/428, 275
PRIORITY FILING DATE: 2003-05-01

PRIOR APPLICATION NUMBER: 09/966545
PRIORITY FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 09/544511
PRIORITY FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: 60/128514
PRIORITY FILING DATE: 1999-04-09

PRIOR APPLICATION NUMBER: 09/569269
PRIORITY FILING DATE: 2000-05-11

PRIOR APPLICATION NUMBER: 60/134315
PRIORITY FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 09/619252
PRIORITY FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/789390
PRIORITY FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185548
PRIORITY FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 450
SOFTWARE: Cureseq1st version 0.1

SEQ ID NO 139

LENGTH: 2432

TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: CDS
LOCATION: (112) .. (1869)

US-10-428-275-139

Query Match 94.0%; Score 1643.8; DB 8; Length 2432;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 CTCGAGACCATGGAGAGGGGACGCCACGGGAATGCAATCTCCACCAAGAAACCTTCAAC 60
Db 102 CTCGAGACCATGGAGAGGGGACGCCACGGGAATGCAATCTCCACCAAGAAACCTTCAAC 161
Qy 61 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 120
Db 162 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 221
Qy 121 CCAGGCACTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 180
Db 222 CCAGGCACTCCAGCCAGGATCTCCAGCTGGGACACCTCCAGCCGGGATCTCCAGC 281
Qy 181 CCAGGCACTCCAGCTGGTACCTCCAGGCGGGGACATCTCCAGCCG----- 228
Db 282 CCAGGCACTCCAGCTGGTACCTCCAGGCGGGGACATCTCCAGCCGCGGATCTCCAGC 341
Qy 229 ---GGCATCTCCAGCCAGGATCTCCAGCCGGGACATCTCCAGCTGTGACATCTTC 285
Db 342 CCAGGCACTCCAGCCAGGATCTCCAGCCGGGACATCTCCAGCTGTGACATCTTC 401
Qy 286 CAGGTCTCATCCGGCAGGTATATCCGCGAGGTCAAGCTCGGTGACAACTTCCCAAC 345
Db 402 CAGGTCTCATCCGGCAGGTATATCCGCGAGGTCAAGCTCGGTGACAACTTCCCAAC 461
Qy 346 CAGGTGATCCTGTTAGAGCAACAGAGTGGGGGCTGTACCCATCCGATCATCTCTGC 405
Db 462 CAGGTGATCCTGTTAGAGCAACAGAGTGGGGGCTGTACCCATCCGATCATCTCTGC 521
Qy 406 CAGGTGACACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 465
Db 522 CAGGTGACACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 581
Qy 466 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 525
Db 582 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 641
Qy 526 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 585
Db 642 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 701
Qy 586 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 645
Db 702 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 761
Qy 646 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 705
Db 762 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 821
Qy 706 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 765
Db 822 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 881
Qy 766 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 825
Db 882 CAGGTGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 941
Qy 826 TGGCCACAGGATTTTCCCAACAGCTTCTCAATTTAGATACAACTCCACATCCAGGA 885
Db 942 TGGCCACAGGATTTTCCCAACAGCTTCTCAATTTAGATACAACTCCACATCCAGGA 1001
Qy 886 AAGGCTCCACAGGCTGAATGCCCTTCCAGCGGATATCTCCCTCCAGTGTCCACTG 945
Db 1002 AAGGCTCCACAGGCTGAATGCCCTTCCAGCGGATATCTCCCTCCAGTGTCCACTG 1061
Qy 946 CGGACTGAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1005
Db 1062 CGGACTGAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1121
Qy 1006 GCTTGGCAAGTGAAGTGTGCACTTCCGACCAACCAATCTGTGAGGACAGCTCATTTGA 1065

Db 1122 GCCTTGCAAGTAGTGTGCACTTGGCAGCACCCACATCTGTGAGGACGCTCATTTGA 1181
Qy 1066 CGCCCATGTGGTGTCTACCTGCCCCCACTCTTTGTGTACCCGGGAGAAAGTCCCTGA 1125
Db 1182 CGCCCATGTGGTGTCTACCTGCCCCCACTCTTTGTGTACCCGGGAGAAAGTCCCTGA 1241
Qy 1126 GGGCTGGAAGTGTACCGGGGACACGAACTTGCACAGTTGGCTGAGGACCTCCAT 1185
Db 1242 GGGCTGGAAGTGTACCGGGGACACGAACTTGCACAGTTGGCTGAGGACCTCCAT 1301
Qy 1186 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGAGAGCACTATGACATGCGCTT 1245
Db 1302 TGCCGAGATCATCATCAACAGCAATTAACCGATGAGAGAGCACTATGACATGCGCTT 1361
Qy 1246 CATGGCTGTCCAAAGCCCTGACCTTGTGCTCAATCCACCTTGTCTCCCTCCAT 1305
Db 1362 CATGGCTGTCCAAAGCCCTGACCTTGTGCTCAATCCACCTTGTCTCCCTCCAT 1421
Qy 1306 GCATGACAGACCTTTAGCCCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGACAG 1365
Db 1422 GCATGACAGACCTTTAGCCCTCAATGAGACCTGTGATCAAGGCTTTGGCAAGACAG 1481
Qy 1366 GGAAGACAGATGACAAAGCATCCCTTCTCTCGGAGGTGAGGTCAATCATCACTT 1425
Db 1482 GGAAGACAGATGACAAAGCATCCCTTCTCTCGGAGGTGAGGTCAATCATCACTT 1541
Qy 1426 CAAGAATGCAATGACTTGTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1485
Db 1542 CAAGAATGCAATGACTTGTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTC 1601
Qy 1486 TGGGAGACCTTGTGGGGGAGAGACTCTGCTGAGGAGAGACAGGGGGGGCTCTGTCTG 1545
Db 1602 TGGGAGACCTTGTGGGGGAGAGACTCTGCTGAGGAGAGACAGGGGGGGCTCTGTCTG 1661
Qy 1546 TGAGCAGAAACCCGCTGTATCTGTGACAGGTGTACAGAGCTGAGGCAAGGCTGTGCA 1605
Db 1662 TGAGCAGAAACCCGCTGTATCTGTGACAGGTGTACAGAGCTGAGGCAAGGCTGTGCA 1721
Qy 1606 GAGAAACAAACCTGTGTGTATACCAAAAGTACAGAAAGTTTCTTCCGTGATTTACAGAA 1665
Db 1722 GAGAAACAAACCTGTGTGTATACCAAAAGTACAGAAAGTTTCTTCCGTGATTTACAGAA 1781
Qy 1666 GATGAG 1672
Db 1782 GATGAG 1788

RESULT 12
US-10-428-275-141
Sequence 141, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428,275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25

NUMBER OF SEQ ID NOS: 450
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 141
LENGTH: 2432
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112) .. (1869)
US-10-428-275-141

Query Match 94.0%; Score 1643.8; DB 8; Length 2432;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 CTGAGAGACCATGAGAGGAGACAGCCAGGAAATGATCTTCCAGACAAACCTTCCAGC 60
Db 102 CTCAGAGACCATGAGAGGAGACAGCCAGGAAATGATCTTCCAGACAAACCTTCCAGC 161
Qy 61 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 120
Db 162 TGGAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 221
Qy 121 CCAAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 180
Db 222 CCAAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 281
Qy 181 CCAAGCATCTCCAGCTGTACCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGC 228
Db 282 CCAAGCATCTCCAGCTGTACCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGC 341
Qy 229 ---GGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 285
Db 342 CCAAGCATCTCCAGCCAGGATCTCCAGCTGGGAGACCTCCAGGCGGCGATCTCCAGC 401
Qy 286 CAGGTCTCATCCGCGAGCTATCATCCGCGAGGTGAGCTTCCAGCTTCCAGC 345
Db 402 CAGGTCTCATCCGCGAGCTATCATCCGCGAGGTGAGCTTCCAGCTTCCAGC 461
Qy 346 CAGGTATCTCTGTATGAGCAACCAAGTGGGGGCTGTATCCATCCGATCATCTCTGC 405
Db 462 CAGGTATCTCTGTATGAGCAACCAAGTGGGGGCTGTATCCATCCGATCATCTCTGC 521
Qy 406 CAGGTACGACCAAGCAACAGGCGACAGGAGAGCCAGGTACGAGCTGCCAAGTT 465
Db 522 CAGGTACGACCAAGCAACAGGCGACAGGAGAGCCAGGTACGAGCTGCCAAGTT 581
Qy 466 CACCTGCGGAGGCGCAGAAAGCTACCGCTCATCGGCTGCTCTCTCATATTGC 525
Db 582 CACCTGCGGAGGCGCAGAAAGCTACCGCTCATCGGCTGCTCTCTCATATTGC 641
Qy 526 CTTGTGTGTGTCTGTATCATCTCTTCCAGTTCTGCGAGGCGCAACAGGATCAGGTA 585
Db 642 CTTGTGTGTGTCTGTATCATCTCTTCCAGTTCTGCGAGGCGCAACAGGATCAGGTA 701
Qy 586 CAAGGACCAAGGAGAGCTGTCCCAAGCAAGCTTGTGAGAGGAGGAGGAGCTG 645
Db 702 CAAGGACCAAGGAGAGCTGTCCCAAGCAAGCTTGTGAGAGGAGGAGGAGCTG 761
Qy 646 CAAGCTGAAGAGTACAGAGCTGGGCTGAGAGGTTGACTGGGACAAAGTCTGCTTAA 705
Db 762 CAAGCTGAAGAGTACAGAGCTGGGCTGAGAGGTTGACTGGGACAAAGTCTGCTTAA 821
Qy 706 AATCTACTGTGGTCTCTCCATCATGAGGCTTCCATCTGTAGACAGAACTGGAATGACTC 765
Db 822 AATCTACTGTGGTCTCTCCATCATGAGGCTTCCATCTGTAGACAGAACTGGAATGACTC 881
Qy 766 CTACTCAGAGAAAGCTGCGAGAGCTGGGTTTGAAGTGTCTACCGGACAAAGGAGT 825
Db 882 CTACTCAGAGAAAGCTGCGAGAGCTGGGTTTGAAGTGTCTACCGGACAAAGGAGT 941
Qy 826 TGCCCAAGGAGATTTTCCAAAGCTTCTCAATCTTGAATACAACTCCACATCCAGGA 885

Db	762	CAACGTGAAGAGTACGACGAGCTGGGCTGGGTGAGGTTTGACTGGGACAAGATCTTGTCTTAA	821
Qy	706	AATCTACTCTTGGGTCTCTCCCATCACTGAGTCTTCCCATCTGTATGACGCAACTGGAAATGACTC	765
Db	822	AATCTACTCTTGGGTCTCTCCCATCACTGAGTGGCTTCCCATCTGTATGACGCAACTGGAAATGACTC	881
Qy	766	CTACTCAGAGAAGACCTGCCAGCAGCTGGGTTTGAAGAGTCTCACCGGAAACCGAAGT	825
Db	882	CTACTCAGAGAAGACCTGCCAGCAGCTGGGTTTGAAGAGTCTCACCGGAAACCGAAGT	941
Qy	826	TGCCCAAGGGATTTTGGCAACAGCTTCTCAATCTTGATATCAACTTCACCATCCAGGA	885
Db	942	TGCCCAAGGGATTTTGGCAACAGCTTCTCAATCTTGATATCAACTTCACCATCCAGGA	1001
Qy	886	AAGCTCTCACAGGTCTGAATGCCCTTCCCAACGGTAATCTCCCTCAGTGTTCACACTG	945
Db	1002	AAGCTCTCACAGGTCTGAATGCCCTTCCCAACGGTAATCTCCCTCAGTGTTCACACTG	1061
Qy	946	CGGACTGAGGGCCATGACCGGGCGGATCGTGGAGGGGCGCTGGCTTCGTATGACCAATG	1005
Db	1062	CGGACTGAGGGCCATGACCGGGCGGATCGTGGAGGGGCGCTGGCTTCGTATGACCAATG	1121
Qy	1006	GCTTTGGCAAGTGAAGTCTGCACTTTCGGCACCACCACTCTGTGAGGCGACGCTCATTTGA	1065
Db	1122	GCTTTGGCAAGTGAAGTCTGCACTTTCGGCACCACCACTCTGTGAGGCGACGCTCATTTGA	1181
Qy	1066	CGCCCAATGGGTGTCACTGCGCGCCCACTGCTTCTTGATGACCGGGAGAAAGTCTGGGA	1125
Db	1182	CGCCCAATGGGTGTCACTGCGCGCCCACTGCTTCTTGATGACCGGGAGAAAGTCTGGGA	1241
Qy	1126	GGGCTGGAAGGTGTACCGGGGCGACAGCAACTGCACCAAGTTGGCTTGAGGCGAGCTTCAT	1185
Db	1242	GGGCTGGAAGGTGTACCGGGGCGACAGCAACTGCACCAAGTTGGCTTGAGGCGAGCTTCAT	1301
Qy	1186	TGCCGAGATCATCATCAACAGCAATTACACCGATGAGGAGGACGACTGTGACATCGCCTT	1245
Db	1302	TGCCGAGATCATCATCAACAGCAATTACACCGATGAGGAGGACGACTGTGACATCGCCTT	1361
Qy	1246	CATCGGGGTGTCCAAAGCCCTGACCCCTGTCCGCTCACATCCACCTGCTTCCCTCCCAT	1305
Db	1362	CATCGGGGTGTCCAAAGCCCTGACCCCTGTCCGCTCACATCCACCTGCTTCCCTCCCAT	1421
Qy	1306	GCATGGAACAGACTTTTACCTCAATGAGACCTGCTGATCAACAGCTTTGGCAAGACGAG	1365
Db	1422	GCATGGAACAGACTTTTACCTCAATGAGACCTGCTGATCAACAGCTTTGGCAAGACGAG	1481
Qy	1366	GGAGAAGATGACAAAGACATCCCTCTTCTCCGGAGGTGAGAGTCAATCTCATATGACTT	1425
Db	1482	GGAGAAGATGACAAAGACATCCCTCTTCTCCGGAGGTGAGAGTCAATCTCATATGACTT	1541
Qy	1426	CAAGAATATGCAATGACTACTTGTGTATGACAGTTACTTACCCCAAGAGATGATGTGTGC	1485
Db	1542	CAAGAATATGCAATGACTACTTGTGTATGACAGTTACTTACCCCAAGAGATGATGTGTGC	1601
Qy	1486	TGGGGACCTTGTGGGGGCGAGACCTCTGCGCAGGGAGACAGCGGGGGGCTCTTGTCTG	1545
Db	1602	TGGGGACCTTGTGGGGGCGAGACCTCTGCGCAGGGAGACAGCGGGGGGCTCTTGTCTG	1661
Qy	1546	TGAGCAGAAACAACGCTGTGTACTGTGGCAGGTGTACCAAGCTGGGGGCAACAGCTGTGGCCA	1605
Db	1662	TGAGCAGAAACAACGCTGTGTACTGTGGCAGGTGTACCAAGCTGGGGGCAACAGCTGTGGCCA	1721
Qy	1606	GAGAAACAACCTGTGTGTGACCAACAAGTGACAGAAGTCTTCCCTTGATTTACAGCA	1665
Db	1722	GAGAAACAACCTGTGTGTGACCAACAAGTGACAGAAGTCTTCCCTTGATTTACAGCA	1781
Qy	1666	GATGGAG 1672	
Db	1782	GATGGAG 1788	

```

: Sequence 123, Application US/10428275
: Publication No. US20040067505A1
:
: GENERAL INFORMATION:
: APPLICANT: Alvarez et al.
: TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
: FILE REFERENCE: 21402-585
: CURRENT APPLICATION NUMBER: US/10/428,275
: CURRENT FILING DATE: 2003-05-01
: PRIOR APPLICATION NUMBER: 09/966545
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/544511
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128514
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 09/569269
: PRIOR FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: 60/134315
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 09/619252
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185548
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 450
: SOFTWARE: Curaseq1ist version 0.1
: SEQ ID NO 123
:
: LENGTH: 1771
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (14) .. (1771)
: US-10-428-275-123

```

RESULT 14
US-10-428-275-123

Query Match	Similarity	93.8%	Score 1640	DB 8	Length 1771
Best Local Match	99.1%	Pred. No. 0			
Matches 1665	Conservative	0	Mismatches	0	Indels 15; Gaps 1
QY	8	ACCATGAGAGGAGACAGCCACGGGAMTGCATCTCCAGCAGAAACACTTCAGCTGGAGCA	67		
Db	11	ACCATGAGAGGAGACAGCCACGGGATGCATCTCCAGCAGAAACACTTCAGCTGGAGCA	70		
QY	68	TCTCCAGGCCGAGGCATCTTCAGCTTGGACACTCTCCAGCCGGGGCATCTCCAGCCGAGCA	127		
Db	71	TCTCCAGGCCGAGGCATCTTCAGCTTGGACACTCTCCAGCCGGGGCATCTTCAGCCGAGCA	130		
QY	128	TCTCCAGGCCGAGGCATCTTCAGCTTGGACACTCCGGGCCGGGGCATCTTCAGCCGAGCA	187		
Db	131	TCTCCAGGCCGAGGCATCTTCAGCTTGGACACTCCGGGCCGGGGCATCTTCAGCCGAGCA	190		
QY	188	TCTCCAGCTGGTACACTTCAGGCCGGGGCATCTTCAGCCCG-----GGCA	232		
Db	191	TCTCCAGCTGGTACACTTCAGGCCGGGGCATCTTCAGCCCGGGGCATCTTCAGCCGAGCA	250		
QY	233	TCTCCAGGCCGAGGCATCTTCAGCCGGGGCATCTCCGGCTCTGGCATCTTCAGGCTCC	292		
Db	251	TCTCCAGGCCGAGGCATCTTCAGCCGGGGCATCTCCGGCTCTGGCATCTTCAGGCTCC	310		
QY	293	TCATCCGGCAGGTCATCATCCGCCAGGTCAGGCTTGATGACAACTCCCAACAGAGTG	352		
Db	311	TCATCCGGCAGGTCATCATCCGCCAGGTCAGGCTTGATGACAACTCCCAACAGAGTG	370		
QY	353	TACCTTTGTAAGCAACACGATGGGGCTTAACCATCCGATCATCTCTCCGAGGTCA	412		
Db	371	TACCTTTGTAAGCAACACGATGGGGCTTAACCATCCGATCATCTCTCCGAGGTCA	430		
QY	413	GCACGAGCAACAGGGCCACGAGGAGAGCCAGGTCAGAGCTTCCCAAGTTTCACTTGG	472		
Db	431	GCACGAGCAACAGGGCCACGAGGAGAGCCAGGTCAGAGCTTCCCAAGTTTCACTTGG	490		
QY	473	CGGAGGGGCCAAGACAGCTACCGTTCATGGGATCGTCTCTCTCATTTGCCCTGGTG	532		

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2006, 08:45:25 ; Search time 1950 Seconds
(without alignments)
1572.094 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 1748
Sequence: 1 ctccagagaccatggagaggg.....ggtcgtcgtgaccgcgagaa 1748

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2330496 seqs, 876882855 residues

Total number of hits satisfying chosen parameters: 4660992

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:
1: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US10_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US11_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US11_NEW_PUB.seq:1*
8: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US11_NEW_PUB.seq:2*
9: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US11_NEW_PUB.seq:3*
10: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1670.4	95.6	2393	US-11-266-748A-22522	Sequence 22522, A
2	709.8	40.6	930	US-11-266-748A-172692	Sequence 172692, A
3	240.2	13.7	255	US-11-266-748A-11696	Sequence 11696, A
4	240.2	13.7	255	US-11-266-748A-64397	Sequence 64397, A
5	240.2	13.7	255	US-11-266-748A-67229	Sequence 67229, A
6	193.2	11.1	2063	US-10-196-749-329	Sequence 329, App
7	193.2	11.1	2063	US-11-101-316-111	Sequence 111, App
8	193.2	11.1	2063	US-11-376-673-111	Sequence 111, App
9	193.2	11.1	2104	US-11-145-307A-34	Sequence 34, App
10	182.8	10.5	2412	US-10-196-749-63	Sequence 63, App
11	181.2	10.4	2413	US-11-400-825-1	Sequence 1, Appl
12	178.4	10.2	2916	US-10-560-723-98	Sequence 98, Appl
13	178.4	10.2	3003	US-10-560-723-96	Sequence 96, Appl
14	178.4	10.2	3102	US-10-560-723-97	Sequence 97, Appl
15	178.4	10.2	3106	US-10-560-723-95	Sequence 95, Appl
16	178.4	10.2	3195	US-11-266-748A-27757	Sequence 27757, A
17	178.4	10.2	3226	US-10-560-723-94	Sequence 94, Appl
18	161.4	9.2	2003	US-11-266-748A-357278	Sequence 357278, A
19	161.4	9.2	2003	US-11-266-748A-440657	Sequence 440657, A
20	154.2	8.8	2544	US-11-400-825-3	Sequence 3, Appl
21	150.4	8.6	2810	US-11-266-748A-31444	Sequence 31444, A
22	148.4	8.5	1713	US-11-266-748A-253384	Sequence 253384, A
23	148.4	8.5	1713	US-11-266-748A-276391	Sequence 276391, A

C	24	148.4	8.5	1713	8	US-11-266-748A-313901	Sequence 313901, A
	25	144.6	8.3	2175	8	US-11-293-697-559	Sequence 559, App
	26	143	8.2	1596	6	US-10-532-668-1	Sequence 1, Appl
	27	135.4	7.7	1199	8	US-11-266-748A-57061	Sequence 57061, A
	28	135.4	7.7	1199	8	US-11-266-748A-62066	Sequence 62066, A
	29	134.4	7.7	639	8	US-11-266-748A-47172	Sequence 47172, A
	30	134.4	7.7	639	8	US-11-266-748A-209672	Sequence 209672, A
	31	133.4	7.6	2788	6	US-10-553-436-98	Sequence 98, Appl
	32	128.4	7.3	786	9	US-11-359-554-1	Sequence 1, Appl
	33	123.6	7.1	847	8	US-11-266-748A-226483	Sequence 226483, A
	34	123.4	7.1	856	8	US-11-266-748A-259941	Sequence 259941, A
	35	123.4	7.1	856	8	US-11-266-748A-320458	Sequence 320458, A
	36	121.6	7.0	1000	8	US-11-266-748A-115746	Sequence 115746, A
	37	121.6	7.0	1000	8	US-11-266-748A-157910	Sequence 157910, A
	38	121.6	7.0	1000	8	US-11-266-748A-224573	Sequence 224573, A
	39	121.6	7.0	1000	8	US-11-266-748A-285377	Sequence 285377, A
	40	121.6	7.0	1000	8	US-11-266-748A-336806	Sequence 336806, A
	41	121.6	7.0	1000	8	US-11-266-748A-395609	Sequence 395609, A
	42	121.6	7.0	1000	8	US-11-266-748A-466655	Sequence 466655, A
	43	120.8	6.9	944	9	US-11-359-554-5	Sequence 5, Appl
	44	120.4	6.9	1122	9	US-11-359-558-1	Sequence 1, Appl
	45	120.4	6.9	1336	8	US-11-266-748A-85241	Sequence 85241, A

ALIGNMENTS

RESULT 1
US-11-266-748A-22522
Sequence 22522, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Hartin, Paul
APPLICANT: Johnston, Patrick
TITLE OF INVENTION: Transcription Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266, 748A
CURRANT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662, 276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700, 293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 48396
SOFTWARE: PatentIn version 3.3
SEQ ID NO 22522
LENGTH: 2393
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-22522
Query Match 95.6%; Score 1670.4; DB 8; Length 2393;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CTCGAGACATGGAGAGGACGACCGGAAATGATCTCCAGCAAGACCTTACG 60
DB 78 CTCGAGACATGGAGAGGACGACCGGAAATGATCTCCAGCAAGACCTTACG 137
OY 61 TGGAGCATCTCCAGCCAGGACATCTCCAGCTGGAGACCTTCAGCGGGCATCTCCAGC 120

138 TGGAGCATCTCCAGCCAGGCGATCTCCAGCTGGAGCAACCTCCAGGCGGCGATCTCCAGC 197
121 CCAGGCACTCTCCAGCCAGGCGATCTCCAGCTGGAGCAACCTCCAGGCGGCGATCTCCAGC 180
198 CCAAGGCACTCTCCAGCCAGGCGATCTCCAGCTGGAGCAACCTCCAGGCGGCGATCTCCAGC 257
181 CCAGGCACTCTCCAGCTGGAGCAACCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGC 240
258 CCAGGCACTCTCCAGCTGGAGCAACCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGC 317
241 CCAGGCACTCTCCAGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 300
318 CCAGGCACTCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 377
301 CAGGTCATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 360
378 CAGGTCATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 437
361 TAGAGCAACACAGTGGGAGCTGTACCCATCCGATCATCTCTCCAGGCGGCGATCTCCAGC 420
438 TAGAGCAACACAGTGGGAGCTGTACCCATCCGATCATCTCTCCAGGCGGCGATCTCCAGC 497
421 AACCAAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
498 AACCAAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 557
481 CCAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
558 CCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 617
541 CATCATCTCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 600
618 CATCATCTCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 677
601 GAGCTGTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 660
678 GAGCTGTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 737
661 CAGCTGTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 720
738 CAGCTGTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 797
721 CTCCATCAGTGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 780
798 CTCCATCAGTGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 857
781 CTCCATCAGTGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 840
858 CTCCATCAGTGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGAT 917
841 TGCCAAAGCTTTCAATCTTGAAGTACAACTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAG 900
918 TGCCAAAGCTTTCAATCTTGAAGTACAACTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAG 977
901 TGAATGCCCTTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 960
978 TGAATGCCCTTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 1037
961 GACCGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
1038 GACCGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1097
1021 TCTGCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1098 TCTGCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1157
1081 CAGTGGCGGCGAGTGGTCTTCTGTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
1158 CAGTGGCGGCGAGTGGTCTTCTGTGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
1141 CGCGGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200

1218 CGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1277
1201 CAACAGCAATTTACACCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
1278 CAACAGCAATTTACACCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1337
1261 GCCCTTGAACCTTGTCCGCTCATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 1320
1338 GCCCTTGAACCTTGTCCGCTCATCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGGCGG 1397
1321 TAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
1398 TAGCTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457
1381 GACATCCCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
1458 GACATCCCTTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1517
1441 CTACTTGTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
1518 CTACTTGTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1577
1501 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
1578 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1637
1561 CTGCTTACCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
1638 CTGCTTACCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1697
1621 TGTGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1672
1698 TGTGTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1749

RESULT 2
US-11-266-748A-172692
Sequence 172692, Application US/11266748A
Publication No. US2006013463A1
GENERAL INFORMATION:
APPLICANT: Harklin, Paul
APPLICANT: Johnston, Patrick
APPLICANT: Mulligan, Karl
TITLE OF INVENTION: Transcriptional Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 172692
LENGTH: 930
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (764)..(764)

OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-112692

Query Match 40.6%; Score 709.8; DB 8; Length 930;
Best Local Similarity 97.5%; Pred. No. 9.8e-155;
Matches 742; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

```
QY 236 CCAGCCAGGATCTCCAGCCCGGCGCATCTCCGCTCGGATCATCTTCAGAGTCCCA 295
DB 1 CCAGCCAGGATCTCCAGCCCGGCGCATCTCCGCTCGGATCATCTTCAGAGTCCCA 60
QY 296 TCCGCGAGGATCATCTCCAGCCAGGTCAGCTCGGTGACAACTCCCAACAGAGTATAC 355
DB 61 TCCGCGAGGATCATCTCCAGCCAGGTCAGCTCGGTGACAACTCCCAACAGAGTATAC 120
QY 356 CTGTGTAGAGCAACACAGTGGGGGCTGTATCCATCCGATATCTCTGCGAGTACCA 415
DB 121 CTGTGTAGAGCAACACAGTGGGGGCTGTATCCATCCGATATCTCTGCGAGTACCA 180
QY 416 CCAGCAACCGAGGCGCACAGGAGAGAGCCAGTACAGCCGAGTCCCAAGTTCACCTGGCG 475
DB 181 CCAGCAACCGAGGCGCACAGGAGAGAGCCAGTACAGCCGAGTCCCAAGTTCACCTGGCG 240
QY 476 GAGGCGCAGAAAGCAGCTACCGCTCATCGGTCGTCCTCTCATTTGCGCTGTGTT 535
DB 241 GAGGCGCAGAAAGCAGCTACCGCTCATCGGTCGTCCTCTCATTTGCGCTGTGTT 300
QY 536 TCGCTCATCATCTCTTTCAGTTCTGCGAGGCGCACAGAGGATCAGTACAGAGCAG 595
DB 301 TCGCTCATCATCTCTTTCAGTTCTGCGAGGCGCACAGAGGATCAGTACAGAGCAG 360
QY 596 AGGAGAGCTGTCCCAAGCAGCTGTCTGTGACCGGGGTGTGAGCTGAACTGAG 655
DB 361 AGGAGAGCTGTCCCAAGCAGCTGTCTGTGAGCGGGGTGTGAGCTGAACTGAG 420
QY 656 AGTACAGAGCTGGGCTGCGTGTGAGCTGGGCAAGTCTGCTTAAATCTACTCT 715
DB 421 AGTACAGAGCTGGGCTGCGTGTGAGCTGGGCAAGTCTGCTTAAATCTACTCT 480
QY 716 GGGTCTCTCCATCAGTGTCTTCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 775
DB 481 GGGTCTCTCCATCAGTGTCTTCCATCTGTAGACAGCACTGGAATGACTCTCACTCAGAG 540
QY 776 AAGACTCTGCGCAGCTGGGTTTGAAGTGTCTCAACCGGAGAACCGAGTGGCCACAG 835
DB 541 AAGACTCTGCGCAGCTGGGTTTGAAGTGTCTCAACCGGAGAACCGAGTGGCCACAG 600
QY 836 GATTTTGCCACAGCTTCTCATCTTGAGATACAACTCCACATCCAGAAAGCTCTCAC 895
DB 601 GATTTTGCCACAGCTTCTCATCTTGAGATACAACTCCACATCCAGAAAGCTCTCAC 660
QY 896 AGGTCTGATGATCCCTTCCAGCGGTATATCT-CCCTCCAGTGTCCCACT-GCGAGCTGA 953
DB 661 AGGTCTGATGATCCCTTCCAGCGGTATATCTCCCTCCAGTGTCCCACTGCGAGCTGA 720
QY 954 GGGCCATGACCGGCGGATCTGTGGAGGGGCGCTGGCTCG 994
DB 721 GGGCCATGACCGGCGGATCTGTGGAGGGGCGCTGG 761
```

RESULT 3

US-11-266-748A-11696/c
Sequence 11696, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnson, Patrick
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 11696
LENGTH: 255
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-11696

Query Match 13.7%; Score 240.2; DB 8; Length 255;
Best Local Similarity 98.8%; Pred. No. 3.4e-46;
Matches 242; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 514 CTCTCTATGCTCCCTGCTGTGTTTGTCTCATCTCTTCCAGTTCTGGCAGGCGCACAC 573
DB 255 CTCTCTATGCTCCCTGCTGTGTTTGTCTCATCTCTTCCAGTTCTGGCAGGCGCACAC 196
QY 574 AGGATAGGTACAGAGAGAGAGAGAGAGCTGTCCCAAGCAGCTGTGGCTGAGCG 633
DB 195 AGGATAGGTACAGAGAGAGAGAGAGCTGTCCCAAGCAGCTGTGGCTGAGCG 136
QY 634 GGTGTGAGCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGGTGAAGTTTGACTGGACAA 693
DB 135 GGTGTGAGCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGGTGAAGTTTGACTGGACAA 76
QY 694 GTCTGTCTTAAATCTACTCTGTGGTCTCCATCACTGCTTCCATCTGTAGACGAA 753
DB 75 GTCTGTCTTAAATCTACTCTGTGGTCTCCATCACTGCTTCCATCTGTAGACGAA 16
QY 754 CTGGA 758
DB 15 CAGCA 11
```

RESULT 4

US-11-266-748A-64397
Sequence 64397, Application US/11266748A
Publication No. US20060134663A1
GENERAL INFORMATION:
APPLICANT: Harkin, Paul
APPLICANT: Johnson, Patrick
TITLE OF INVENTION: Transcriptome Microarray Technology and
FILE REFERENCE: 55815-0102 (319189)
CURRENT APPLICATION NUMBER: US/11/266,748A
CURRENT FILING DATE: 2005-11-03
PRIOR APPLICATION NUMBER: EP 04105479.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105482.6
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105483.4
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2

```

PRIORITY FILING DATE: 2004-11-03
PRIORITY APPLICATION NUMBER: US 60/662,276
PRIORITY FILING DATE: 2005-03-14
PRIORITY APPLICATION NUMBER: US 60/700,299
PRIORITY FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 48396
SOFTWARE: PatentIn version 3.3
SEQ ID NO 64397
LENGTH: 255
TYPE: DNA
ORGANISM: Homo Sapiens
US-11-266-748A-64397

```

Query Match	13.7%	Score 240.2;	DB 8;	Length 255;
Best Local Similarity	98.8%	Pred. No. 3.4e-46;		
Matches 242; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	514	CGCTCTCATTGACCTCGTGTGGTTTTCGCTCATCATCTCTTCCAGTCTGGCAGGGGCACAC	574
Db	1	CGTCTCATTTGCCCTGGTGGTTTCGCTCATCATCTCTTCCAGTCTGGCAGGGGCACAC	60
QY	574	AGGGATTCAGGTTACAAAGGACACAGGGAGAGCTGTCTCCAGCACCGTGTTCCTGTGACCG	633
Db	61	AGGGATTCAGGTTACAAAGGACACAGGGAGAGCTGTCTCCAGCACCGTGTTCCTGTGACCG	120
QY	634	GGTGTGTGACCTGCAAGCTGAAGGTGACAGAGCTGGGCTGCCTGAGGTTTGACTGGGACAA	693
Db	121	GGTGTGTGACCTGCAAGCTGAAGGTGACAGAGCTGGGCTGCCTGAGGTTTGACTGGGACAA	180
QY	694	GTCTCTGCTTTAAATCTTACTCTTGGGTCTCCCATAGTGGCTTCCATCTGTAGACGAA	753
Db	181	GTCTCTGCTTTAAATCTTACTCTTGGGTCTCCCATAGTGGCTTCCATCTGTAGACGAC	240
QY	754	CTGGA	758
Db	241	CAGCA	245

```

1  RESULT 5
2  US-11-266-748A-67229/C
3  Sequence 67229, Application US/11266748A
4  Publication NO. US20060134663A1
5  GENERAL INFORMATION:
6  APPLICANT: Harkin, Paul
7  APPLICANT: Johnston, Patrick
8  APPLICANT: Mulligan, Karl
9  TITLE OF INVENTION: Transcendome Microarray Technology and
10 TITLE OF INVENTION: Methods of Using the Same
11 FILE REFERENCE: 55815-0102 (319189)
12 CURRENT APPLICATION NUMBER: US/11/266,748A
13 CURRENT FILING DATE: 2005-11-03
14 PRIOR APPLICATION NUMBER: EP 04105479.2
15 PRIOR FILING DATE: 2004-11-03
16 PRIOR APPLICATION NUMBER: EP 04105482.6
17 PRIOR FILING DATE: 2004-11-03
18 PRIOR APPLICATION NUMBER: EP 04105483.4
19 PRIOR FILING DATE: 2004-11-03
20 PRIOR APPLICATION NUMBER: EP 04105507.0
21 PRIOR FILING DATE: 2004-11-03
22 PRIOR APPLICATION NUMBER: EP 04105485.9
23 PRIOR FILING DATE: 2004-11-03
24 PRIOR APPLICATION NUMBER: EP 04105484.2
25 PRIOR FILING DATE: 2004-11-03
26 PRIOR APPLICATION NUMBER: US 60/662,216
27 PRIOR FILING DATE: 2005-03-14
28 PRIOR APPLICATION NUMBER: US 60/700,293
29 PRIOR FILING DATE: 2005-07-18
30 NUMBER OF SEQ ID NOS: 48396
31 SOFTWARE: PatentIn version 3.3
32 SEQ ID NO 67229
33 LENGTH: 255
34 TYPE: DNA
35 ORGANISM: Homo Sapiens

```

US-11-266-748A-67229

Query Match	13.7%	Score 240.2;	DB 8;	Length 255;
Best Local Similarity	98.8%;	Pred. No. 3.4e-46;		
Matches 242;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	51	CGTCCTCATATGGCCCTGGTGGTTCGCTCATCATCTCTTCCAGTTCTGGCAGGGCCACAC	578
Db	255	CGTCCTCATATGGCCCTGGTGGTTCGCTCATCATCTCTTCCAGTTCTGGCAGGGCCACAC	198
Qy	574	AGGGATCAGGTACAGGAGCAGAGGGAAGCTGTCTCCACGACGCTGTTCCGCTGTGACGG	633
Db	195	AGGGATCAGGTACAGGAGCAGAGGGAAGCTGTCTCCACGACGCTGTTCCGCTGTGACGG	138
Qy	634	GGTGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGACTGGGACAA	693
Db	135	GGTGTGTGACCTGCAAGCTGAAGAGTGAAGAGCTGGGCTGCGTGAAGTTTGACTGGGACAA	76
Qy	694	GTCTCTGCTTTAAATCTACTCTTGGGCTCTCCCATATAGTGGCTTCCCATCTGTACAGCAA	758
Db	75	GTCTCTGCTTTAAATCTACTCTTGGGCTCTCCCATATAGTGGCTTCCCATCTGTACAGCAA	16
Qy	754	CTGGA 758	
Db	15	CAGCA 11	

```

1 RESULT 6
2 US-10-196-749-329
3 Sequence 329, Application US/10196749
4 Publication No. US20060094864A1
5 GENERAL INFORMATION:
6 APPLICANT: Baker, Kevin P.
7 APPLICANT: Chen, Jian
8 APPLICANT: Deenoyers, Luc
9 APPLICANT: Goddard, Audrey
10 APPLICANT: Godowski, Paul J.
11 APPLICANT: Gurney, Austin L.
12 APPLICANT: Pan, James
13 APPLICANT: Smith, Victoria
14 APPLICANT: Watanabe, Colin K.
15 APPLICANT: Wood, William I.
16 APPLICANT: Zhang, Zhen
17 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
18 FILE REFERENCE: P3430R1C340
19 CURRENT APPLICATION NUMBER: US/10/196,749
20 CURRENT FILING DATE: 2002-07-16
21 PRIOR APPLICATION NUMBER: 10/052566
22 PRIOR FILING DATE: 2002-01-15
23 PRIOR APPLICATION NUMBER: 60/059263
24 PRIOR FILING DATE: 1997-09-18
25 PRIOR APPLICATION NUMBER: 60/059266
26 PRIOR FILING DATE: 1997-09-18
27 PRIOR APPLICATION NUMBER: 60/062250
28 PRIOR FILING DATE: 1997-10-17
29 PRIOR APPLICATION NUMBER: 60/063120
30 PRIOR FILING DATE: 1997-10-24
31 PRIOR APPLICATION NUMBER: 60/063121
32 PRIOR FILING DATE: 1997-10-24
33 PRIOR APPLICATION NUMBER: 60/063486
34 PRIOR FILING DATE: 1997-10-21
35 PRIOR APPLICATION NUMBER: 60/063540
36 PRIOR FILING DATE: 1997-10-28
37 PRIOR APPLICATION NUMBER: 60/063541
38 PRIOR FILING DATE: 1997-10-28
39 PRIOR APPLICATION NUMBER: 60/063544
40 PRIOR FILING DATE: 1997-10-28
41 NUMBER OF SEQ ID NOS: 612
42 SEQ ID NO 329
43 TYPE: DNA
44 LENGTH: 2063
45

```


Qy 1444 TTCTCCCTGGATTATACAGCAAGATGAGACGAGGTGCGATTGCAAAATCTTAACCAAG 1703
Db 1468 ATCTCAACTGGATTACAAATGTCGTGAAGGCTGAGCTTAATGCTGCTGCTTTGACG 1527

RESULT 8
US-11-376-673-111

Sequence 111, Application US/11376673
Publication No. US20060160186A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin I.
APPLICANT: Wood, William I.
TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
TITLE OF INVENTION: LUNG TUMOR
FILE REFERENCE: P3230R1C165C
CURRENT APPLICATION NUMBER: US/11/376,673
CURRENT FILING DATE: 2006-03-14
PRIOR APPLICATION NUMBER: 10/063742
PRIOR FILING DATE: 2002-05-09
PRIOR APPLICATION NUMBER: 10/06867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 111
LENGTH: 2063
TYPE: DNA
ORGANISM: Homo Sapien
US-11-376-673-111

Query Match 11.1%; Score 193.2; DB 9; Length 2063;
Best Local Similarity 55.8%; Pred. No. 5.5e-35;
Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

Qy 924 TCTCCCTCCAGTGTCCCACTGCGGACTGAGGAGCCATCCGCGGCGATCGTGGAGAGGG 983
Db 760 TCTCCCTCCAGTGTCTTCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 819
Qy 984 CGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTGGGCAACCCACA 1043
Db 820 AGGAGGCTCTGTGGATTCTTGGCTTGGCAAGTGAAGTCTGCACTTGGGCAACCCACA 879
Qy 1044 TCTGTGAGGAGCAGCTCATTTAGCCGCCAGTGGTGTCTGCTGCGGCCCACTGCTTTG 1103
Db 880 TCTGTGAGGAGCAGCTCATTTAGCCGCCAGTGGTGTCTGCTGCGGCCCACTGCTTTG 935
Qy 1104 TGACCCGGAGAGAGTCTGAGAGGCTGGAAGGTGTAGCGGGGACCAAGCAACTGACAC 1163
Db 936 --AGGAAACATACCGATGTGTTCATCTGAAGGTGCGGGGAGGCTCAGACAACTGGACA 993
Qy 1164 AGTGGCTGAGGAGCCTTCATTTGCGGAGATCATCATCAACAGCAATTACCGGATGAG 1223
Db 994 GCTTCCC--ATCCGTGGCTGTGGCCAGATCATCATTAATTAACCCCATGTACC 1050
Qy 1224 AGGAGCACTATGACATGCGCTCATGCGGCTGTGTCAGAGCCCTGACCTGTTCCTACA 1283
Db 1051 CCAAGACATATGACATGCGCTCATGCGGCTGTGTCAGAGCCCTGACCTGTTCCTACA 1110
Qy 1284 TCCACCCCTGTGCTGCTCCCATGATGACAGACCTTTAGCTCAATGAGACTGTGGA 1343
Db 1111 TCAGGCCATCTGTGCTGCTCTTTTGTATGAGAGCTCATCTCAGACCCCACTTGGGA 1170
Qy 1344 TCACAGGCTTTGGCAAGACAGGAGACAGATGACAAAGACATTCCTCTCCGCGAGG 1403
Db 1171 TCATTGATGGGGCTTTTACAGAGGATGAGAGGGAATATGTCTGACATCTGCTGACG 1230
Qy 1404 TGACAGTCAATCTCATGCACTTTCAAGAAATGCAATGACTTGTGTTATATACGTTACC 1463

Db 1231 CGTCAGTCCAGTCAATTGACAGACACAGGTGCAATGACAGATGCTTACCAGGGGAGAG 1290
Qy 1464 TTACCCCAAGATGATGTGTGTGGAGCACTTGTGGGGGCAAGAGATCTCTCCAGGAG 1523
Db 1291 TCACCGAAGATGATGTGTGTGGAGCACTTGTGGGGGCAAGAGATCTCTCCAGGAG 1350
Qy 1524 ACAGCGGGGGCTCTTGTCTGTGAGCAAGAACACCGCTGTATCTGAGGTTCACCA 1583
Db 1351 ACAGTGTGGGGCCCTGATGT--ACCAATCTGACAGTGGCATGTGTGGGATCTGTTA 1407
Qy 1584 GCTGGGGCAGAGGTGTGGCCAGAGAAACAACTGTGTGTATACCAAGTGCACAGAG 1643
Db 1408 GCTGGGGCTATGTGTGGGGGGCCCGAGCACCCAGAGATATACCAAGGTCTTACCT 1467
Qy 1444 TTCTCCCTGGATTATACAGCAAGATGAGAGCGAGTGCATTCAGAAAATCTTAACCA 1703
Db 1468 ATCTCAACTGGATTACAAATGTCTGGAAGGCTGATGTATGCTGCTGCCCTTTGACG 1527

RESULT 9
US-11-145-307A-34

Sequence 34, Application US/11145307A
Publication No. US20060094035A1
GENERAL INFORMATION:
APPLICANT: Arcuturus Bioscience, Inc.
APPLICANT: Erlander, Mark G.
APPLICANT: Ma, Xiao-Jun
TITLE OF INVENTION: Identification of Tumors
FILE REFERENCE: 022041-002020US
CURRENT APPLICATION NUMBER: US/11/145,307A
CURRENT FILING DATE: 2005-06-03
PRIOR APPLICATION NUMBER: US 60/577,084
PRIOR FILING DATE: 2004-06-04
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34
LENGTH: 2104
TYPE: DNA
ORGANISM: Homo sapiens
US-11-145-307A-34

Query Match 11.1%; Score 193.2; DB 8; Length 2104;
Best Local Similarity 55.8%; Pred. No. 5.6e-35;
Matches 435; Conservative 0; Mismatches 333; Indels 12; Gaps 3;

Qy 924 TCTCCCTCCAGTGTCCCACTGCGGACTGAGGAGCCATCCGCGGCGATCGTGGAGAGGG 983
Db 798 TCTCCCTCCAGTGTCTTCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
Qy 984 CGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACTTGGGCAACCCACA 1043
Db 858 AGGAGGCTCTGTGGATTCTTGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 917
Qy 1044 TCTGTGAGGAGCAGCTCATTTAGCCGCCAGTGGTGTCTGCTGCGGCCCACTGCTTTG 1103
Db 918 TCTGTGAGGAGCAGCTCATTTAGCCGCCAGTGGTGTCTGCTGCGGCCCACTGCTTTG 973
Qy 1104 TGACCCGGAGAGAGTCTGAGAGGCTGGAAGGTGTAGCGGGGACCAAGCAACTGACAC 1163
Db 974 --AGGAAACATACCGATGTGTTCATCTGAAGGTGCGGGGAGGCTCAGACAACTGGACA 1031
Qy 1164 AGTGGCTGAGGAGCCTTCATTTGCGGAGATCATCATCAACAGCAATTACCGGATGAG 1223
Db 1032 GCTTCCC--ATCCGTGGCTGTGGCCAGATCATCATTAATTAACCCCATGTACC 1088
Qy 1224 AGGAGCACTATGACATGCGCTCATGCGGCTGTGTCAGAGCCCTGACCTGTTCCTACA 1283
Db 1089 CCAAGACATATGACATGCGCTCATGAGAGCTGCAAGTTCACATTCATCTTCCAGGACAG 1148
Qy 1284 TCCACCCCTGTGCTGCTCCCATGATGACAGACCTTTAGCTCAATGAGACTGTGGA 1343
Db 1149 TCAGGCCATCTGTGCTGCTCTTTTGTATGAGAGCTCATCTCAGACCCCACTTGGGA 1208

Query 1344 TCACAGCCTTTGGCAGAGCCAGAGACAGATGACAAAGATCCCTCTCCGGGAG 1403
Db 1209 TCATTGGATGGGCTTTACAGACAGATGAGAGGAAATCTCGACATATCTGCTGACAG 1268
Qy 1404 TGCAGTTCATCTCATGCTTCAAGAAATGCAATGACTACTGCTGTATGACAGTTACC 1463
Db 1269 CGTGAATTCAGATGATGACAGACACAGGTCATATGACAGATGCGTACAGAGGGAG 1328
Qy 1464 TTACCCAGAGATATGATGCTGAGGACCTTCTGGGGGAGAGACTCTCCAGAGAG 1523
Db 1329 TCACCGAAGAAATATATGTCAGAGCATCCCGAAGGGGATGAGACCTGCGCAGGCTG 1388
Qy 1524 ACACGGGGGAGCTCTTGTCTGTGAGCAGAACACCGCTGTATCTGACAGGTGACCA 1583
Db 1389 ACAGTGTGGGCTCTGATGT--ACCAATCTGACCAATGAGTGTGAGGATCTGTTA 1445
Qy 1584 GCTGGGGCAGAGCTGTGGCCAGAGAAACAACTGTGTGTACACCAAGTGAAGAG 1643
Db 1446 GCTGGGGCTATGCTGGGGGGCCCGAGACCCAGAGGATATACACCAAGGTCTCAGCCT 1505
Qy 1644 TTCTCCCTGATTTACAGCAAGATGAGAGCGAGTCCGATTCAGAAATCTTAACAG 1703
Db 1506 ATCTCAATGATCTACAAATGTCTGGAAGGCTGAGCTGTATGCTGCTGCCCTTTGAG 1565

RESULT 10
US-10-196-749-63
Sequence 63, Application US/10196749
Publication No. US2006094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Macanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 63
LENGTH: 2412
TYPE: DNA
ORGANISM: Homo Sapien
US-10-196-749-63

Query Match 10.5%; Score 182.8; DB 6; Length 2412;
Best Local Similarity 55.8%; Pred. No. 1,56-33;
Matches 415; Conservative 0; Mismatches 317; Indels 12; Gaps 3;
Qy 953 AGGGCCATGACCGGGCGGATCTGGAGAGGGCCGTGCTCGATATGACAGTGGCTTGG 1012
Db 756 AGGGGCTACAGCTCAGCCATCTGGTGGTGAAGATGTCCTTGTCTCGCAGTGGCCCTGG 815
Qy 1013 CAAGTGAATCTGCACTTGGGACCCACCAATCTGTGAGGACGCTCATTTAGAGCCAG 1072
Db 816 CAGGCCAGCCCTTCAGTTCAGAGGCTACCACTGTGCGGGGGCTCTGTATCACGCCCCG 875
Qy 1073 TGGGTGCTCACTGCGCGCCCACTGCTCTTGTGTGACCGGGAGAGGCTCGAGGGCTGG 1132
Db 876 TGGATCATCACTGCTGACACTGTGTTATG-----ACTTTGATCTCCCAAGTCAATG 929
Qy 1133 AAGGTGTACGCGGGCACCAACCTGACCACTGCTGAGGAGCCCTCCATTGCGAG 1192
Db 930 ACCATCAGGTGGGTCTAGTTTCCCTTGGAGCAATCAGCCCATCCATTGTTGAG 989
Qy 1193 ATCA---TCATCAACAGCAATTACACCGATGAGAGAGCACTTGTGATGCTGCTCATG 1249
Db 990 AAGATTGCTTACACACAGCAAGTACAAAGCAAGAGGCTGGGCAATGACATGCCCTTATG 1049
Qy 1250 CGGCTGTCAAGCCCGCTGACCCCTGCTCCGCTCACATCCACCTGCTGCTCCCATGAT 1309
Db 1050 AAGCTGCGCGGGCCACTCAGCTTCAATGAATGAATGTCAGGCTGTGCTCCCACTCT 1109
Qy 1310 GGAACAGACTTTAGACCTCAATGAGACCTGTGATGATCAGAGCTTTGGCAGAGCAGGAG 1369
Db 1110 GAAGAGAACTTCCCGCATGAGAAAGTGTGCTGACGTCAGAGATGGGGGGCCAGAGAGAT 1169
Qy 1370 ACAGATGACAAAGATCTCCCTTCTCCGGAGGTGCAAGTCAATCTCATGCACTTCAAG 1429
Db 1170 GAGAGTAC--GCTCTCCCTGTCTCAACACACCGGCGCTCTTTGATTTCCAAAG 1226
Qy 1430 AAATGCAATGACTTGTGTCTATGACAGTTAATCTTACCCCAAGAGATGATGCTGGG 1489
Db 1227 ATCTGCAACCAAGAGAGGTATGAGGTGAGCATCTCCCTTCAGCTCTGAGGGGCG 1286
Qy 1490 GACCTTGTGGGGGACAGACTCTGTCAGAGAGACAGCGGGGGCCCTTGTCTGTGAG 1549
Db 1287 TACCTGACGGGTGGCTGTGACAGCTGCAAGGGGAGACAGCGGGGGCCCTGTGTGTA 1346
Qy 1550 CAGAAACACCGCTGTACTGTGACAGGTGTCAACAGCTGGGGCAGAGCTGTGCGCAGAG 1609
Db 1347 GAGAGAGGCTGTGAGAGTTAGTGGAGCGACAGCTTTGGCATCGGCTGCGCAGAGGTG 1406
Qy 1610 AACCAACCTGTGTGTACACCAAGTGACAGAGTCTTCCCTGTGATTTACAGCAAGTG 1669
Db 1407 AACCAAGCTGGGGGTACACCGGTGTCACTCTTCTGTGATGTACAGAGCAAGTG 1466
Qy 1670 GAGAGCAGGTGCGATTCAAGAAA 1693
Db 1467 GAGAGAGACTTAAACCTGAAGA 1490

RESULT 11
US-11-400-825-1
Sequence 1, Application US/11400825
Publication No. US20060177866A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
FILE REFERENCE: D6192C1P3
CURRENT APPLICATION NUMBER: US/11/400,825
CURRENT FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: 10/455,720
PRIOR FILING DATE: 2003-06-05
NUMBER OF SEQ ID NOS: 154

```
; SEQ ID NO 1
; LENGTH: 2413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; OTHER INFORMATION: entire cDNA sequence of TADG-12 gene
US-11-400-825-1
```

```
Query Match      10.4%; Score 181.2; DB 7; Length 2413;
Best Local Similarity 55.6%; Pred. No. 3.5e-32;
Matches 414; Conservative 0; Mismatches 318; Indels 12; Gaps 3;
```

```
QY 953 AGGGCCATGACCGGGCGGATGTTGGAGAGGGCGCTGGCCCTCGAGTAGCAGTGGCTTGG 1012
Db 777 AGGGGCTACAGCTACCGCATGTTGGTGAAGAAATGTCCTTGTCTGCGACGTGGCCCTGG 836
QY 1013 CAATGAGTCTGCACTTGGGACCAACCAATCTGTGGAGGCAAGCTCATTTGACGCCAG 1072
Db 837 CAGGCGACCTTCAGTTCAGAGGCTACCACTGTGCGGGGCTCTGTATCATCACGCCCTTG 896
QY 1073 TGGGCTGCTACGCGCCGACCTGCTTCTTGTGACCCGGAGAAAGTCTGGAAGGGCTGG 1132
Db 897 TGAATCATCTACTGCTGACACACTGTGTTATG-----ACTTGTACTCTCCCAAGTCATGG 950
QY 1133 AAGGTGTACGCGGGACCAAGCAACCTGCAACAGTTGCTTGAAGGCAAGCTTCATTTGCCAG 1192
Db 951 ACCATTCAGAGTGGGTCTAGTTTCCCTGTGTGACATCCAGCCCATTCCTCATTTGGTGGAG 1010
QY 1193 ATCA---TCATCAACAGCAATTTACACGATAGAGAGACATATGATCATGCTCCCTCATG 1249
Db 1011 AAGATTGTCTACACAGCAAGTATCAAGCCAAAGAGGCTGGGCAATGATCATGCTCTTATG 1070
QY 1250 CGGCTGTCAAGCCCTGACCCCTGACCTGCTGACATCCACCTGCTGCTGCTCCCATGAT 1309
Db 1071 AAGCTGGCCGGGACCACTACGTTCAATAAATGATCCAGCTGTGTGCTTCCCAACTCT 1130
QY 1310 GGACAGACCTTTAAGCTTCATATAGACCTGTGTATCACAGGCTTTGGCAAGACAGGAG 1369
Db 1131 GAAGAGAACTTCCCGATGAGAAAGTGTGTGACGTGACGATGGGGGGCCACAGAGGAT 1190
QY 1370 ACAATGACAAAGACATCCCCCTCTCTCCGGAGGTGACGTCAATCTCATGACTTCAAG 1429
Db 1191 GGAGGTGAC---GCTCCCTGTGTCTGTAACCAACAGCGGCGCTCCCTTGTATTTCCAAAG 1247
QY 1430 AAATGCAATGACTACTTGATCTATAGACATTACCTTACCCCAAGATGATGTGTGGG 1489
Db 1248 ATCTGCAACCAAGAGGACGTGTACGTTGCGTCAATCTCCCTCCATGCTCTGGCGGGGC 1307
QY 1490 GACCTTCTGCGGGGACAGACCTCCCTGACGAGGAGACAGCGGGGGCTCTTGTCTGTAG 1549
Db 1308 TACCTGACCGGCTGGGTGAACAGCTGCCAGGGGGACAGCGGGGGCCCTGTGTGTCAA 1367
QY 1550 CAGAACACCGCTGTACTCTGACAGGTGTACACAGCTGGGGCAACAGGCTGTGGCCAGAG 1609
Db 1368 GAGAGGAGGCTGTGAGATTAGTGGAGCGACCAAGCTTGTGCATCTGCTGGCGCAGAGGTG 1427
QY 1610 AACCAACCTGTGTGTATACCAAGTGAAGAGTCTTCCCTGATTTACAGCAAGATG 1669
Db 1428 AACCAAGCTGGGGGTGTACACCGGTGTACCTCTTCTGAGCTGATCCAGACAGCATGTG 1487
QY 1670 GAGAGCGAGGTGCAATTCAGAAAA 1693
Db 1488 GAGAGAGACCTTAAAAAAGCTGAAGA 1511
```

```
RESULT 12
US-10-560-723-98
; Sequence 98, Application US/10560723
; Publication No. US20060115821A1
; GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTRA
```

```
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 80213WO
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 2916
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-98
```

```
Query Match      10.2%; Score 178.4; DB 6; Length 2916;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;
```

```
QY 969 GGATCTGGAGAGGGCGCTGCGCTCGATAGCAAGTGGCTTGGCAAGTGTCTGCACT 1028
Db 581 GGATCTGGAGGGGTGAGAGCGCGCTCCCGGGGCTGGCCCTTGGCAGGTCAAGCTTGCAAG 640
QY 1029 TCGGACACACCATCTGTGTGAGACAGGCTCATTTAGCGCCAGTGGGTGCTCACTGCGG 1088
Db 641 TCCGAACGTCACAGTGTGCGAGGCTTCATCATCACCCCGAATGGATCTGTACAGCGG 700
QY 1089 CCACTGCTTCTTGTGACCCGGAGAGAGTCTCGAAGGCTGGAAGTGTACCGGGGCA 1148
Db 701 CCACTGCTGTGAAAAAAGCTTTAAATCATGATGGAACGGCATTTGGCGGGGATTT 760
QY 1149 CCAGCAACTGACACAGTGTGCTTGAAGCGCTTCATTTGCCAGAG---TCATCATCAACA 1205
Db 761 TGACACATCTTCTATGTATGTATGAGCCGAGTACCAAGTAGAAAAAGTGTATCTATC 820
QY 1206 GCAATTTACCCGATGAGAGAGACGATATGACATGCGCCCTCATGCGGCTGTCCAAAGCCC 1265
Db 821 CAATTTATGACTCCAAAGACCAAGAACATGACATTTGGCTGTATGAAGTGTGAGAGGCTC 880
QY 1266 TGACCTGTGCGCTCACATCCACCTGCTTGCCTCCCATGATGACAGACCTTTAGCC 1335
Db 881 TGACTTTCAAGCACTGTGTGATTTCCGGGTGGGGGACACCGAGAGAG---AAGGAAAGCT 940
QY 1326 TCAATGAGACCTGTGTATACAGGCTTTGGCAAGACCAAGGAGACAGATGTACAAGACAT 1385
Db 941 CAGAACAGCTCTGTGTGATTTCCGGGTGGGGGACACCGAGAGAG---AAGGAAAGCT 997
QY 1386 CCCCCTTCTCCGGGAGGTGAGCTCAATCTCATGCACTTCAAGAAATGCAATGACTACT 1445
Db 998 CAGAAAGTGTGAAGCGTCCAGAGGTGCTTCTCATTTGAGACAGAGATGTCAACAGACAT 1057
QY 1446 TGATCTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGACCTTGTGTGGGGCA 1505
Db 1058 ATGTCTATGACCACTGATACACCAAGCATGATCTGTGCGGCTTCTGTGAGGGGAAAG 1117
QY 1506 GAGACTCTGCGCAGAGGAGACAGCGGGGGCTCTTGTCTGTGAGCAGAAACACCGCTGT 1565
Db 1118 TCGATTCTTTCAGAGGAGACAGTGTGAGGGGCTCTGTGTCACTTCAAGAAACAAATCTGT 1177
QY 1566 ACCGTGAGAGGTGTACCAAGCTGGGGCAACAGCTGTGGCAGAGAAACAACTGTGTGT 1625
Db 1178 GGTGTATAGGGGATACAAAGCTGGGGTTCTGTCTGTGCAAAAGCTTACAGACAGAGATGT 1237
QY 1626 ACACCAAGTGTACGAAGTCTTCTCCCTGATTTACAGCAAGATGAGAGCGAGGTGCAAT 1685
Db 1238 ACCGGAATGTGATGTATTCACGGACTGGATTTATGCAAAATGAGGAGAGACGGCTAAT 1297
QY 1686 TCAGAAATCTT 1697
Db 1298 CCACATGTGCTT 1309
```

```
RESULT 13
US-10-560-723-96
; Sequence 96, Application US/10560723
; Publication No. US20060115821A1
```

GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
; FILE REFERENCE: B0213WO
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-96

Query Match 10.2%; Score 178.4; DB 6; Length 3003;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGGGCGCTGGCCCTCGATACAGTGGCCCTTGGCAAGTGAAGTCTGCACT 1028
DB 668 GGATCGTGGAGGGGCGCTGGCCCTCGATACAGTGGCCCTTGGCAAGTGAAGTCTGCACT 727
QY 1029 TCGGCACCAACCCATCTGTGAGAGCAAGCTCATTTGACGCCAGTGGGTCTCACTGGCG 1088
DB 728 TCCAGAAAGTCTCAAGTGTGGAGAGGCTTCATCATCACCCTCGAGTGAATCGTACAGCCG 787
QY 1089 CCCACTGCTTCTTGCTGAACCCGGGAGAAAGTCTGAGAGGCTGGAAGGTGATACGCGGCA 1148
DB 788 CCCACTGCTGAGAAAACCTTTAAACAATTCATGGCATTTGACGCGCATTTGCGGGGATTT 847
QY 1149 CCAGCAACTGACCAAGTGGCTTGAAGGAGCTTCATTGCGGAGA--TCAATCATCAACA 1205
DB 848 TGAACAATCTTTATGTTTATGAGAGCGGATACCAAGTAAAGTATTTCTCATC 907
QY 1206 GCAATTCACCGATGAGAGAGACGACTATGATCGCCCTCATGGCGGCTGCCAAGCCCC 1265
DB 908 CAAATTTATGATCCAAAGCAAGAAACAATGACATTTGGCTATGAGTGCAGAGGCTTC 967
QY 1266 TGACCTGTGCGCTCAACATCCACCTGCTGCTCCCTCCCATGATGACAGACCTTTAGCC 1325
DB 968 TGACTTTCAAGACCTAGTAAACCAAGTGTCTGCCCAACCGAGCATGATGCTGCAGC 1027
QY 1326 TCAATGAGACTGCTGATCAACAGGCTTTGGCAAGACGAGGAGACAGATGACAGACAT 1385
DB 1028 CAGAACAGCTCTGCTGGAATTTCCGGGTGGGGGCCACCGAGAGAG--AAGGGAAGACT 1084
QY 1386 CCCCCCTTCGCGGAGGTGAGAGTCAATCTCATGACTTTCAGAAATGCAATGACTACT 1445
DB 1085 CAGAAAGTGTAAAGCTGCCAAGGTTCTCATTTGAGACACAGAGATGCAACAGCAT 1144
QY 1446 TGCTCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGACCTTGTGGGGCA 1505
DB 1145 ATGCTCTATGACACCTGATCAACACAGCATGATGTGTGCGGCTTCTGCGAGGGAACG 1204
QY 1506 GAGACTCTGCGCAGGAGACACAGCGGGGCTCTTGTCTGTGAGCAAAACCGCTGGT 1565
DB 1205 TCGATTTCTGCGAGGAGTACAGTGGAGGCTCTGTGCTCACTTCGAAAGAACATATCTGT 1264
QY 1566 ACCTGGCAGGTGTCAACGAGCTGGGGCAAGGCTGGGCGCAGAGAAACAACTGGTGTGT 1625
DB 1265 GCGTGTATGGGAGATACAGCTGGGTTCTGGCTGTGCCAAGCTTACAGACCGAGAGTGT 1324
QY 1626 ACACCAAGTACAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCAT 1685
DB 1325 ACGGGAATGTATGTATTTACAGCACTGGAATTTATCGAACAAATAGAGGACAGCGCTAAT 1384
QY 1686 TCAGAAATCTCT 1697
DB 1385 CCACATGTCTTT 1396

RESULT 14

US-10-560-723-97
; Sequence 97; Application US/10560723
; Publication No. us20060115821A1
; GENERAL INFORMATION:
; APPLICANT: EXONHIT THERAPEUTICALS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
; FILE REFERENCE: B0213WO
; CURRENT APPLICATION NUMBER: US/10/560,723
; CURRENT FILING DATE: 2005-12-15
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-560-723-97

Query Match 10.2%; Score 178.4; DB 6; Length 3102;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGGGCGCTGGCCCTCGATACAGTGGCCCTTGGCAAGTGAAGTCTGCACT 1028
DB 767 GGATCGTGGAGGGGCGCTGGCCCTCGATACAGTGGCCCTTGGCAAGTGAAGTCTGCACT 826
QY 1029 TCGGCACCAACCCATCTGTGAGAGCAAGCTCATTTGACGCCAGTGGGTCTCACTGGCG 1088
DB 827 TCCAGAAAGTCTCAAGTGTGGAGAGGCTTCATCATCACCCTCGAGTGAATCGTACAGCCG 886
QY 1089 CCCACTGCTTCTTGCTGAACCCGGGAGAAAGTCTGAGAGGCTGGAAGGTGATACGCGGCA 1148
DB 887 CCCACTGCTGAGAAAACCTTTAAACAATTCATGGCATTTGACGCGCATTTGCGGGGATTT 946
QY 1149 CCAGCAACTGACCAAGTGGCTTGAAGGAGCTTCATTGCGGAGA--TCAATCATCAACA 1205
DB 947 TGAACAATCTTTATGTTTATGAGAGCGGATACCAAGTAAAGTATTTCTCATC 1006
QY 1206 GCAATTCACCGATGAGAGAGACGACTATGATCGCCCTCATGCGGCTGTGCCAAGCCCC 1265
DB 1007 CAAATTTATGATCCAAAGCAAGAAACAATGACATTTGGCTATGAGTGCAGAGGCTTC 1066
QY 1266 TGACCTGTGCGCTCAACATCCACCTGCTGCTCCCTCCCATGATGACAGACCTTTAGCC 1325
DB 1067 TGACTTTCAAGACCTAGTAAACCAAGTGTCTGCCCAACCGAGCATGATGCTGCAGC 1126
QY 1326 TCAATGAGACTGCTGATCAACAGGCTTTGGCAAGACCGAGGAGACAGATGACAGACAT 1385
DB 1127 CAGAACAGCTCTGCTGGAATTTCCGGGTGGGGGCCACCGAGAGAG--AAGGGAAGACT 1183
QY 1386 CCCCCCTTCGCGGAGGTGAGAGTCAATCTCATGACTTTCAGAAATGCAATGACTACT 1445
DB 1184 CAGAAAGTGTAAAGCTGCCAAGGTTCTCATTTGAGACACAGAGATGCAACAGCAT 1243
QY 1446 TGCTCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGACCTTGTGGGGCA 1505
DB 1244 ATGCTCTATGACACCTGATCAACACAGCATGATGTGTGCGGCTTCTGCGAGGGAACG 1303
QY 1506 GAGACTCTGCGCAGGAGACACAGCGGGGCTCTTGTCTGTGAGCAAAACCGCTGGT 1565
DB 1304 TCGATTTCTGCGAGGAGTACAGTGGAGGCTCTGTGCTCACTTCGAAAGAACATATCTGT 1363
QY 1566 ACCTGGCAGGTGTCAACGAGCTGGGGCACAGGCTGTGGCAGAGAAACAACTGGTGTGT 1625
DB 1364 GCGTGTATGGGAGATACAGCTGGGTTCTGCTGTGCCAAGCTTACAGACCGAGAGTGT 1423
QY 1626 ACACCAAGTACAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGCGAGTGCAT 1685
DB 1424 ACGGGAATGTATGTATTTACAGCACTGGAATTTATCGAACAAATAGAGGACAGCGCTAAT 1483
QY 1686 TCAGAAATCTCT 1697
DB 1484 CCACATGTCTTT 1495

RESULT 15
US-10-560-723-95
/ Sequence 95: Application US/10560723
/ Publication No. US20060115821A1
/ GENERAL INFORMATION:
/ APPLICANT: EXONHIT THERAPEUTICALS
/ TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF AS TARGETS FOR PROSTA
/ TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
/ FILE REFERENCE: B0213M0
/ CURRENT APPLICATION NUMBER: US/10/560.723
/ CURRENT FILING DATE: 2005-12-15
/ NUMBER OF SEQ ID NOS: 185
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 95
/ LENGTH: 3106
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-560-723-95

Query Match 10.2%; Score 178.4; DB 6; Length 3106;
Best Local Similarity 54.6%; Pred. No. 1.7e-31;
Matches 400; Conservative 0; Mismatches 326; Indels 6; Gaps 2;

QY 969 GGATCGTGGAGGAGGCGCTGGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCACT 1028
DB 771 GGATCGTGGGCGGTGAGAGCGCGCTCCCGGGGCGCTGGCCTGGCAGGTCAAGCTGCAAG 830
QY 1029 TCGGCACCAACCATCTGTGAGAGGCAAGCTCATTTGACGCCAGTGGGTGCTCACTGCCG 1088
DB 831 TCCAGAAAGTCCAGTGTGCGAGGAGCTCCATCATCACCCCGAGTGAATGCTGACAGCCG 890
QY 1089 CCCACTGCTTCTTGTGTGATCCCGGAGAAAGTCTGSAAGGCTGGAAGTGTACCGGAGCA 1148
DB 891 CCCACTGCGTGAAGAAAGCTCTTAAACAATCATGCAATTTGACCGCATTTGCGGGGATTT 950
QY 1149 CCAGCAACTGACCAAGTGTGCTGAGGAGCGCTCCATTTGCCGAGA---TCATCATCAACA 1205
DB 951 TGAGACATATCTTTCATGTCTTATGAGCGGATACCAAGTAGAAAGATTTCTCATC 1010
QY 1206 GCAATTACACCGATGAGAGAGACATATGATATCGCCCTCATAGCGCTGTCCAAAGCCCC 1265
DB 1011 CAAATTATGATCTCAAGACCAAGAACATGATTCGCTGATGAGCTGCAAGGCTTC 1070
QY 1266 TGACCTGTCCGCTCAATCCACCTGCTTGCCTCCCATGATGACAGACCTTTAGCC 1325
DB 1071 TGACTTTCAAGACCTTATGTAACCAAGTGTCTGCCCAACCAAGCATGATGCTGCAGC 1130
QY 1326 TCAATGAGACTGCTGTGATCAAGGCTTTGGCAAGACCAAGGAGACAGATGACAAGACAT 1385
DB 1131 CAGAACAGCTCTGCTGATTTCCGGGTGGGGGCGCACCGAGGAGA---AAGGGAAGACT 1187
QY 1386 CCCCCCTTCCTCGGAGAGGTGACATCTCATGCACTTCAAGAAATGCAATGACTACT 1445
DB 1188 CAGAAAGTGTGACGCTGCCAAGGTGCTTCTCATTTGAGACACAGAGATGCAACAGCAGAT 1247
QY 1446 TGGTCTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGAGCTTGTGTGGGAGCA 1505
DB 1248 ATGCTATGACCACTGATCAACACAGCATGATCTGTGCCGCTTCTTGCAGGGGAAAG 1307
QY 1506 GAGACTCTGTCAGAGGAGACAGCGGGGGGCTTGTGTGTGAGCAAGAACACCGCTGGT 1565
DB 1308 TCGATTTCTTGCAGAGGTGACAGTGAAGGCGCTCTGTGTCACTTCGAAGAACATATCTGGT 1367
QY 1566 ACCGTGAGAGGTGTACCAAGCTGGGGGCAAGGCTGTGGCCAGAGAAACAACTGTGTGT 1625
DB 1368 GCGTGTATAGGGGATTAAGGCTGGGCTTCTGCTGTGTCCAAAGCTTTACAGACAGAGAGTGT 1427
QY 1626 ACAACCAAGTGTACAGAAATTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGGTGCAT 1685
DB 1428 ACGGGAATGTGATGTATTCACGGAATGATTTTATTCAGCAAAATGAGGCGAGACCGCTAAT 1487

QY 1686 TCAGAAAATCCT 1697
DB 1488 CCACATGTGCTTT 1499

Search completed: September 18, 2006, 13:15:32
Job time : 1952 secs

PR	24-MAY-2002	2002US-0387830P
PR	24-MAY-2002	2002US-0387830P
PR	24-MAY-2002	2002US-0384044P
PR	30-MAY-2002	2002US-0384215P
PR	30-MAY-2002	2002US-0384315P
PR	30-MAY-2002	2002US-0384366P
PR	30-MAY-2002	2002US-0384397P
PR	30-MAY-2002	2002US-0384432P
PR	30-MAY-2002	2002US-0384452P
PR	31-MAY-2002	2002US-0385411P
PR	02-JUL-2002	2002US-0393333P
PR	09-AUG-2002	2002US-0402154P
PR	09-AUG-2002	2002US-0402171P
PR	09-AUG-2002	2002US-0402204P
PR	09-AUG-2002	2002US-0402205P
PR	22-AUG-2002	2002US-0405175P
PR	23-AUG-2002	2002US-0406129P
PR	23-SEP-2002	2002US-0412954P
PR	30-SEP-2002	2002US-0414497P
PR	07-OCT-2002	2002US-0416661P
PR	24-OCT-2002	2002US-0420851P
PR	31-OCT-2002	2002US-0422547P
PR	01-MAY-2003	2003US-00428275

PA (CURA-) CURAGEN CORP

PI Alvarés E, Anderson DM, Boldog FL, Cameron E, Edinger SR,
PI Fernandes ER, Gerlach VL, Gorman L, Grose WM, Guo X, Ji W,
PI Kekuda R, Li L, MacDougall JR, Padigara M, Patursajan M,
PI Peterson JD, Raselli L, Shimmels RA, Splyek KA, Stone DJ,
PI Vernet CM, Voss EZ, Zhong M,

DR WPI; 2004-053040/05.

DR N-PSDB; ADH17439.

PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections

PS Claim 1; SEQ ID NO 130; 478bp; English.

CC The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC candiant, hypotensive, antihypertensioleptotic, anorectic, virucide,
CC antihypertensive, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteoprotic, antistatic,
CC antinflammatory, dermatological, antiaesthetic and antilipemic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorder, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.

Sequence 569 AA;

Alignment Scores:

Pred. No.:	4,19e-181	Length:	569
Score:	3004.00	Matches:	5630
Percent Similarity:	100.00	Conservative:	0
Best Local Similarity:	100.00	Mismatches:	0
Query Match:	90.1%	Indels:	0
DB:	8	Gaps:	0

US-10-806-370-11 (1-1748) X ADH17440 (1-569)

OY	8	CCCAATGGAGAGGGACACGACCGGGAAATGCATCTCCACAGAAACAACCTTACAGCTGGAGCA	67
Db	4	ThnrtGtUatgAapSerHtAsgLYaShAlaSerProAlaArgHtrProSerAlaGtLYaA	23
OY	68	TCCTCAGCCCGAGGATCTCCAGCTGGGAGACAACCTCCAGGCGGGGATCTCCAGCCAGGCA	127
Db	24	SerProAlaGlnAlaSerProAlaGtLYThrProProGtLYArgAlaSerProAlaGlnAla	43
OY	128	TCCTCAGCCCGAGGATCTCCAGCTGGGAGACAACCTCCGGGCGGGGATCTCCAGCCAGGCA	187
Db	44	SerProAlaGlnAlaSerProAlaGtLYThrProProGtLYArgAlaSerProAlaGlnAla	63
OY	188	TCCTCAGCTGGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCAGGCA	247
Db	64	SerProAlaGtLYThrProProGtLYArgAlaSerProGtLYArgAlaSerProAlaGlnAla	83
OY	248	TCCTCAGCCCGGAGCATCTCCGGCTCTGGGCATCACTTTCAGGTCCTCATCCGGAGGTC	307
Db	84	SerProAlaArgAlaSerProAlaLeuAlaSerLeuSerAlaArgSerSerGtLYArgSer	103
OY	308	TCATCCGCGAGGTACAGCTCGGTGACACAACCTCCCAACCGAGAGTACCTGTATAGACA	367
Db	104	SerSerAlaArgSerAlaSerValTThrThSerProThrArgValTYrLeuValAlaArgAla	123
OY	368	ACACCAAGTGAGGCTGTAACCCATCGATTCATCTCTGCGAGGTACAGCACGACCAACGAG	427
Db	124	ThrProValGtLYaValaValProLeArgSerSerProAlaArgSerAlaProAlaThrArg	143
OY	428	GGCAACGAGGAGACCCAGAGTACGAGGCTGCGCAAGTTCACCTGGCGGGAGGGCAGAG	487
Db	144	AlaThrArgLeuSerProGtLYThrSerLeuProArgPheHtrThrArgGtLYGlnIly	163
OY	488	CAGCTACCGCTCATCGAGTGGGTGCTCTCTCTCATTCGCCGTGATGTTGGCTCATATC	547
Db	164	GlnLeuProLeuIleGtLYCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIle	183
OY	548	CTCTTCAGATTTGGGAGGGGACACACAGGATACAGTACAGAAAGACAGGAAAGCTGT	607
Db	184	LeuHegGlnPheTrpGlnGtLYHtAsthGtLYLeArgTYrLysGtLYGlnArgGlnSerCys	203
OY	608	CCCAAGACAGCTGTCCGCTGACGAGGAGTGTGACATGCAAGCTGAAGAGTACAGACTG	667
Db	204	ProLYHtAlaValaArgCysAaArgLYaValaAaArgCysLYuSerLYuSerAaArgLYu	223
OY	668	GGCTGCGGTAGGTTTAACTGGAGCAAGTCTGCTTAAATCTACTCTGGTCTCCCAT	727
Db	224	GtLYaValaArgPheAaPheArgLYuSerLeuSerLYuSerGtLYuSerSerHtA	243
OY	728	CAGTGGCTCCCATCTGTAGAGAGCAACTGGAAATGACTCTTACAGAAAGACTGCGCAG	787
Db	244	GlnTrpLeuProIleCysSerSerHtTrpAaAaPheArgTYrSerGtLYuThCysGln	263
OY	788	CAGCTGGGTTTGGAGTGGCTCACCGGACAAACGAGGTTGGCCACAGGATTTTGGCAAC	847
Db	264	GlnLeuGtLYrHegLYuSerAlaHtAaArgThrThGtLYaAlaHtAaArgPheAlaAaH	283
OY	848	AGCTTTCATCTTGAATACAAATCCACACATCCAGAAAGGCTCCACAGGCTCGAATGC	907
Db	284	SerPheSerIleLeuAaArgTYrAaSerThrIleGlnIuSerLeuHtAaArgSerGtLYu	303
OY	908	CCTTCCCAAGCGTATATCTCCCTCCATGTATCCCAATGCGGATGAGAGGCGCATGACCGG	967
Db	304	ProSerGlnArgTYrIleSerLeuGlnCysSerHtCysGtLYuAaArgAlaMetThrGtLY	323
OY	968	CGAATCGTGGAGGGGCGCTGGCTCCGATAGCAAGTGGCTTGGCAATGATGCTGCAC	1027
Db	324	ArgIleValaGtLYaAlaLeuAlaSerAaPheSerLYeTrpProTrpGlnAlaSerLeuHtA	343
OY	1028	TTCCGACACCAACCATCTGAGAGGACCGCTATGACCGCCAGTGGGTCTCATGCGC	1087

Db	344	PhegilythrThrIhrlHisIleIscysGlyGlyIYhrlLeuIleAspAlaIleGlnTrpValLeuThrIa	365
QY	1088	GGCCACGTCGTTCTTCGTGACCCCGGAGAAAGTCTGTGAGGGCTGGAACTGTACGGCGGC	1147
Db	364	AlahIscsPhePheValThrArgGluIuIysValIleuGluGlyTrpIysValIYhrlaGly	383
QY	1148	ACGAGCAACCGGACCGAGTGGCTGAGGACGCTCCATTGGCCGATGATCATCAACGC	1207
Db	384	ThrsrAsnLeuHISGlnLeuProGlnAlaIleSerIleAlaGlnIleIleIleAsnSer	403
QY	1208	AATATACACCGATGAGAGAGAGCACTATGACATCGCCCTCATGCGGCTGTCCAAAGCCCTG	1267
Db	404	AsnIYhrlhAspGluGluIuAspAspIYhrlAspIleAlaLeuMetArgIleuSerIYhrlProIeu	423
QY	1268	ACCCTGTCCGCTCAATCCACCTGCTGTTGCTCCCTCCCATGATGACAGACCTTTAGCCTC	1327
Db	424	ThrIeuSerAlaHisIleHisIleProIaCysIeuProMetHisGlyGlnThrPheSerIeu	443
QY	1328	AATGAGACCTCTCTGGATCAGCGCTTTGGCAAGACACGAGGAGACAGATGACAAACATCC	1387
Db	444	AsnIuIhrlhCysTrpIleThrlGlyPheGlyIYhrlArgGluThrAspAspIYhrlhSer	463
QY	1388	CCCTTCCTCCGGGAGGTGACAGTCAATCTCATCCGATTCAGAAATGCAATGACTACTTG	1447
Db	464	ProPheIeuArgIuValGlnValAsnLeuIleAspPheIuIysCysAsnAspIYhrlIeu	483
QY	1448	GTCATATGACATTACTTACCCCAAGATGATGTGTGCTGGGAGCTTCGTGGGGGCAGA	1507
Db	484	ValIYhrlhAspSerIYhrlIeuThrProArgMetMetCysAlaGlyAspIeuArgGlyIYhrlArg	503
QY	1508	GACTCTCTGCCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAGAACACCGCTGGTAC	1567
Db	504	AspSerCysGlnGlyAspSerGlyIYhrlProIeuValCysGlnGlnAsnAsnArgTrpIYhrl	523
QY	1568	CTGGCAGGTGTCAACAGCTGGGGGACAGGCTGTGGCCAGAAACAACTGGTGTAC	1627
Db	524	LeuAlaGlyValIhrlhSerTrpGlyThrIYcysGlyGlnArgAsnIYhrlhProGlyValIYhrl	543
QY	1628	ACCAAGTGACAGAAATTCTTCCTCGATTTTACAGCAAGATGAGAGCGAGGTGGCATTC	1687
Db	544	ThrlYsValIhrlhGluValIeuProTrpIleIYhrlSerIYhrlMetGluSerGluValArgPhe	563
QY	1688	AGAAATCC 1696	
Db	564	ArgIYhrlSer 566	
RESULT 2			
AAE17238			
ID	AAE17238	standard; protein; 562 AA.	
XX	AAE17238;		
AC			
XX			
DT	18-APR-2002	(first entry)	
XX			
DE		Human transmembrane serine protease.	
XX			
KW		Human; transmembrane serine protease; gene therapy; metastasis; tumour;	
KW		chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;	
KW		atherosclerosis; neurodegenerative disease; neuroprotective; cyostatic;	
KW		pathogenic infection; antiinflammatory; antiarteriosclerotic;	
KW		antibacterial.	
XX			
OS		Homo sapiens.	
XX			
PN	WO200196538-A2.		
XX			
PD	20-DEC-2001.		
XX			
PF	12-JUN-2001; 2001WO-EP006618.		
XX			
PR	13-JUN-2000; 2000US-0211224P.		
PR	13-APR-2001; 2001US-0283353P.		
PR	16-APR-2001; 2001US-0283648P.		

XX	(FARB) BAYER AG.
PA	Xiao Y, Gedrich R;
PI	WPI: 2002-098065/13.
XX	M-RSDB; AAD27734.
DR	
XX	
PT	Novel isolated polynucleotide encoding transmembrane serine protease
PR	polypeptide, for treating chronic obstructive pulmonary disease, tumor
PT	angiogenesis, inflammation, atherosclerosis and neurodegenerative
PT	disease.
XX	
PS	Claim 1, Fig 1, 120pp: English.
XX	
CC	The present invention relates to an isolated polynucleotide encoding a
CC	transmembrane serine protease polypeptide. Transmembrane serine protease
CC	gene is useful in gene therapy. The invention also relates to a
CC	pharmaceutical composition which is useful for modulating the activity of
CC	transmembrane serine protease in a disease, such as chronic obstructive
CC	pulmonary disease (COPD), metastasis of malignant cells, tumour
CC	angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease
CC	or pathogenic infection. Transmembrane serine protease is useful as a
CC	bait protein in a two-hybrid or three-hybrid assay. The polypeptide is
CC	useful for generating antibodies against it and in various assay systems
CC	The present sequence is a human transmembrane serine protease
XX	
SQ	Sequence 562 AA:
Alignment Scores:	
Pred. No.:	8,66e-181 Length: 562
Score:	2999.00 Matches: 562
Percent Similarity:	100.0% Conservative: 0
Best Local Similarity:	100.0% Mismatches: 0
Query Match:	90.0% Indels: 0
DB:	Gaps: 0
US-10-806-370-11 (1-1748) x AAEL17238 (1-562)	
OY	11 ATGAGAGGGACACGCCGGAATGCATCTCCAGACAAMACACTTCAGTGTGAGCATCT 70
Dd	1 MeGLInrghspserhisaglysnalaserProAlahgtnrrProSerAlaglyAlaser 20
OY	71 CCA GCCCAGGCATCTTCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCAGGCATCT 130
Dd	21 ProAlaglnlaserProAlaglythrProproglyAgaIaserProAlaglnlaser 40
OY	131 CCA GCCCAGGCATCTTCAGCTGGGACACCTCCAGGCCCGGGCATCTCCAGCCAGGCATCT 190
Dd	41 ProAlaglnlaserProAlaglythrProproglyAgaIaserProAlaglnlaser 60
OY	191 CCA GCTGTAAACCCTCCAGCCGGGCATCTCCAGGCCGGGCATCTCCAGCCAGGCATCT 250
Dd	61 ProAlaglyThrProProglyAgaIaserProGlyAgaIaserProAlaglnlaser 80
OY	251 CCA GCCCGGCATCTCCGGCTCTGGCATCACTTTCAGAGTCTCTGATCCGAGGTCATCA 310
Dd	81 ProAlarAgIaserProAlaleunlaserleuseraIgserserSerGlyArGseSer 1000
OY	311 TCCGCCAGGTCAGCTCGGTGA CAACCTTCCCACACAGAGTGTAACCTTTTAGAGCACAA 370
Dd	101 SerAlarGserlaserValthrSerProItraGvalTYLeuValArgAlaThr 1200
OY	371 CCA GTGGGGGTATACCACTCCGATCACTCTCCGAGTGTGAGACACAGAACAGAGGCC 430
Dd	121 ProValglYAlaValaProlleargserseProAlahgserralaprolalrhnaIarla 1400
OY	431 ACCAGGAGAGGCCAGAGTACGAGCTGCCAATTCACTTGCGGAGGCCAGAGAAG 490
Dd	141 ThrArgIuserProolyThrsierleuProlysPheThrTPArgGluglYglnlyGln 1600
OY	491 CTACCGCTATCGGGTGGTGTCTCTCTCATATGCTCGTGGTGTGGCTCATATCTC 5560

Db 161 LeuProLeu11eg1yCyVal1LeuLeu11eal1eVal1SerLeu11e1Leu 180
 Qy 551 TTCCAGTTCTGGCAGGCGCACAGAGATCAGTACAGAGAGAGAGAGAGTGTCCC 610
 Db 181 PheGlnPheTprGlnGlyH1eThrG1y1LeaArgTyr1ySg1uGln1ArgS1uSerCyero 200
 Qy 611 AAGCAGCGTGTGGCTGTGACGGGGGTGTGACATGCAAGCTGAAAGAGTGAAGTGGCC 670
 Db 201 LyS1sAla1eVal1aCySAspG1yVal1aAspCyS1yS1eU1yS1eAspG1uLeG1y 220
 Qy 671 TGGGTGAGGTGTGACTGGGCAAGCTGTCTTAAATCTACTGGGTCTCCCATCAG 730
 Db 221 CyVal1aArgPheAspTPrAspLyS1eU1yS1eTyrS1eG1yS1eSerH1eGln 240
 Qy 731 TGGCTCCCATCTGTAGCAGCACTGGAATGACTCTACTAGAGAAAGCTGGCAGAG 790
 Db 241 TrpLeuPro11eCyS1eSerSerAsnTPrAspAspSerTyrS1eG1u1yS1eThCyG1uGln 260
 Qy 791 CTGGGTTTCCAGAGTCTCACCCGACAAACGAGGTGGCCACAGGATTTTCCACAGC 850
 Db 261 LeuG1yPheG1uSerAlaH1eArgThrG1uVal1aH1eSArgAspPheAlaAsnSer 280
 Qy 851 TTTCATCTTGAAT1aCAACTCCACCATCCAGAAAGCTCCACAGGTCTGAATGCCCT 910
 Db 281 PheSer11eU1aArgTyrAsnSerThr11eGln1uSer1eU1aSArgS1eG1uCyep1ro 300
 Qy 911 TCCAGCGGTATATCTCCCTCAGTGTTCCTCAGTCCGAGTGAAGGCGCATGCCGAG 970
 Db 301 SerGlnArgTyr11eSer1eU1nCyS1eSerH1eCySg1yLeu1aArg1aMetThrG1yArg 320
 Qy 971 ATCGTGGAGGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCATT 1030
 Db 321 11eVal1eG1yG1yAla1eU1a1aSerAspSerTyrTPrG1n1Val1eSer1eU1a1aSpHe 340
 Qy 1031 GGCACACCCCATCTGTGAGAGCAGCTCATTAACGCCGAGTGGGTCTACTGCCGCC 1090
 Db 341 G1yThrThrH1eS1e1eCySg1yG1yThr1eU1eAsp1aG1nTPrVal1eThrH1a1a1a 360
 Qy 1091 CACTGCTTCTTCTGTGACCCCGAGAGAGGTCTGAGAGGCTGAAAGTGTACGCGGAC 1150
 Db 361 H1eCyS1ePhePheVal1ThrArgG1u1yVal1eU1nG1yTPr1yS1eVal1yAlaG1yThr 380
 Qy 1151 AGCAACTGCAACCGTGTGCTGAGCAGCTTCCATGGCCGATCATCATCAACGCAAT 1210
 Db 381 SerAsn1eU1aSg1n1eU1aProG1uAla1aSer11eAlaG1u1e1e1e1eAsnSerAsn 400
 Qy 1211 TACACCGATGAGAGAGCACTATGACATGCGCCCTCAGGCTGTCCAAAGCCCTGAC 1270
 Db 401 TyrThrAspG1uG1uAspAspTyrAsp11eAla1eU1eMetArg1eU1eSer1eU1eThr 420
 Qy 1271 CTGTCCGCTCATCAACCCCTGCTTGCCTCCCATGATGACAGACCTTATGCTCAAT 1330
 Db 421 LeuSerAlaH1eS1e1eH1eProAla1Cy1eU1ePro1eMetH1eSg1yG1n1ThrPheSer1eU1aSn 440
 Qy 1331 GAGACCTGCTGGATCAACGGCTTGGCAACACGAGGAGACATGACACATCCCC 1390
 Db 441 G1uThrCyS1eTPr1eH1nG1yPheG1yLyS1eThrArgG1uThrAspAspLyS1eThSerPro 460
 Qy 1391 TTCTCCGGGAGGAGGATCAATCTCATGACTTCAAGAAATGAATGACTATGCTGTC 1450
 Db 461 Phe1eU1aArg1uVal1aGlnVal1aAsn1eU1eAspPhe1yS1yCySAspAspTyr1eU1a1 480
 Qy 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTCTGTGGGACCTTCTGTGGGCGACAGAC 1510
 Db 481 TyrAspSerTyr1eU1nThrProArg1eMetCyS1aG1yAsp1eU1aArgG1yG1yArgAsp 500
 Qy 1511 TCTGCGCAGGAGACAGCGGGGGGCTTGTCTGTGAGCAAGAAACACCGCTGGTACTG 1570
 Db 501 SerCySg1nG1yAspSerG1yG1yPro1eU1a1CySg1uG1n1aAsnArgTPrTyr1eU1 520
 Qy 1571 GCAGGTGACCACTGGGCGACAGGCTGTGGCCAGAGAAACAAACCTGGGTGTACACC 1630
 Db 521 AlaG1yVal1ThrSerTPrG1yThrG1yCySg1yG1n1aArgAsnLyS1eProG1yVal1yThr 540

Qy 1631 AAAGTGAAGAAGTTCTTCCCTGGATTACAGCAAGATGAGAGAGCGAGTGCATTACA 1690
 Db 541 LyVal1ThrG1uVal1eU1aProTPr11eTyrS1eTyrMetG1uSerG1uVal1aArgPheArg 560
 Qy 1691 AAATCC 1696
 Db 561 LysSer 562
 RESULT 3
 AAB01943
 ID AAE01943 standard; protein: 562 AA.
 AC AAE01943;
 AC AAE01943;
 DT 31-JUL-2001 (first entry)
 XX Human transmembrane serine protease (Endothelinase 2-8) protein.
 DE Human; endothelinase 2-8; protease domain; cytosolic; vulnary; wound;
 KW nocrotic; peridontitis; dermatological disorder; gene therapy; scar;
 KW angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
 KW chronic inflammatory disease; ocular disorder; circulatory disorder;
 KW crest syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
 KW liver cirrhosis; osteoradionecrosis; systemic sclerosis; osteopagial;
 KW inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
 KW systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
 KW transmembrane serine protease.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FH 321. 562
 FT Domain /label= Protease_domain
 FT
 PN MO200136604-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000MO-UE031803.
 XX
 PR 18-NOV-1999; 99US-0166391P.
 PR 22-SEP-2000; 2000US-0234840P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Madison BL, Ong BO;
 XX
 DR WPI; 2001-336001/35.
 DR N-PSDB; AAD05796.
 XX
 PT New nucleic acid encoding a protein comprising endothelinase activity,
 PT useful in the prevention and treatment of e.g. vascular malformations,
 PT cardiovascular disorders, and chronic inflammatory disease.
 PS
 PS Claim 4; Page 138-139; 152pp; English.
 XX
 CC The present sequence is human short form transmembrane serine protease
 CC (endothelinase 2-8) protein. Endothelinase 2 is a type-II membrane-type
 CC serine protease which has a transmembrane domain at the N-terminus,
 CC followed by a single low density lipoprotein-A receptor domain and a
 CC single scavenger-receptor cysteine-rich domain. The C-terminus of the
 CC endothelinase 2 contains the trypsin-like serine protease domain
 CC characterised by the catalytic triad residues in 3 highly conserved
 CC regions of the catalytic domain. In addition 3 repetitive sequence are
 CC found just before the transmembrane domain and represent a sequence motif
 CC for N-myristoylation modification. The invention relates to an
 CC endothelinase protein, endothelinase protease domain and their
 CC corresponding nucleic acid molecules. An endothelinase protein or protease
 CC domain of it is useful for the treatment and diagnosis of disorders
 CC associated with aberrant angiogenesis or undesired neovascularisation.
 CC The undesired angiogenesis is associated with disorders selected from
 CC solid neoplasm, vascular malformations and cardiovascular disorders such

CC as angiofibroma, angiolipoma, atherosclerosis, restenosis/reperfusion
CC injury, arteriovenous malformations, haemangiomatosis and vascular
CC adenoma, dyschondroplasia with vascular hamartomas (Pafucelli's
CC syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
CC syndrome) and Von Hippel Lindau syndrome, chronic inflammatory bowel
CC disease, diabetes mellitus, haemophilic joints, inflammatory bowel
CC disease, nonhealing fractures, periodontitis, psoriasis, rheumatoid
CC arthritis, venous stasis ulcers, granulations-burns, hypertrophic scars,
CC liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
CC granuloma and systemic scleritis and aberrant wound repairs, circulatory
CC disorders Raynaud's phenomenon, cret syndromes such as calcinosis,
CC oesophageal, dyomeolysis, sclerodactyly and teangiectasis, dermatological
CC disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
CC stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
CC caused by ocular neovascular disease, corneal graft neovascularization,
CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia
CC and corneal neovascularisation. The nucleic acids of the invention are
CC also used in gene therapy. The invention also provides method for
CC screening compounds that modulate angiogenesis

XX Sequence 562 AA;

Alignment Scores:
Pred. No.: 2.77e-180 Length: 562
Score: 2991.00 Matches: 561
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 89.7% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x AAE01943 (1-562)

QY 11 ATGGAGAGGAGCAGCAGCGGATGATCTTCACAGAAAGACCTTCAGTGCATCT 70
Db 1 MetGlnArgAspSerHisGlyAsnHisSerProHisArgThrProSerHisGlyAlaSer 20
QY 71 CCAGCCGAGGATCTCCAGCTGGAGACCTTCAGGCGGCGATCTCCAGCCGAGCATCT 130
Db 21 ProHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 40
QY 131 CCAGCCGAGGATCTCCAGCTGGAGACCTTCAGGCGGCGGATCTCCAGCCGAGCATCT 190
Db 41 ProHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 60
QY 191 CCAGCTGGTACCTCCAGGCGGCGATCTCCAGGCGGCGATCTCCAGCCGAGCATCT 250
Db 61 ProHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 80
QY 251 CCAGCCGAGGATCTCCAGCTGGAGACCTTCAGGCGGCGATCTCCAGGCGGCGATCTCA 310
Db 81 ProHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 100
QY 311 TCAGCCGAGGATCTCCAGCTGGAGACCTTCAGGCGGCGATCTCCAGGCGGCGATCTCA 370
Db 101 SerHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 120
QY 371 CCAGTGGGGGTGTAACCATCCGATCTCTTCAGTGCATCTTCAGGCGGCGGCGGCGGCGG 430
Db 121 ProHisGlnHisSerProHisGlyThrProProGlyArgAlaSerProHisGlnHisSer 140
QY 431 ACCAGGAGAGCCAGGTAACAGCTTCGCAAGTTCACTTCGCGGCGGCGGCGGCGGCGGAG 490
Db 141 ThrArgGlnSerProGlyThrSerLeuProHisSerProHisGlnHisSerProHisGln 160
QY 491 CTACCGCTCATCGGGTGGTCTCTCTCATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 550
Db 161 LeuProLeuHisGlyCysValLeuLeuLeuLeuValSerLeuLeuLeuLeu 180
QY 551 TTCAGTTCCTGGCAGGCGCAGCAGGATCAGGTACAAGAGAGAGAGAGAGAGAGAGAGAG 610
Db 181 PheGlnPheTrpGlnGlyHisThrGlyLeuArgTyrLeuGlnGlnArgGlnSerCysPro 200

QY 611 AACGACGCTGTGCTGTGACGGGGGTGTGACCTGCAGCTGAAGATGACGAGCTGGGC 670
Db 201 LysHisHisAlaValArgCysAspGlyValAlaAspCysLeuLeuSerHisArgGlnLeuGly 220
QY 671 TGGGTGAGGTTTACTGGAGCAAGCTCTGCTTAAATCTACTCTGGGCTCTCCATCAG 730
Db 221 CysValArgPheAspTrpAspLysSerLeuLeuLeuLeuLeuSerGlySerHisGln 240
QY 731 TGGCTTCCCATCTGTACAGCAACTGGAATGATCTCTTACAGAGAGAGAGAGAGAGAG 790
Db 241 TrpLeuProLeuLeuSerHisArgTrpAspSerTrpSerGlyTrpTrpCysGlnGln 260
QY 791 CTGGGTTTGGAGGTGTCACCGGACAGGAGGTTGCCACAGGATTTTGGCAAGCAGC 850
Db 261 LeuGlyPheLeuSerHisAlaHisArgTrpTrpHisGlnAlaHisArgAspPheHisSer 280
QY 851 TTCTCATCTTGAGATCAACCTCCACCATCCAGAAAGCTTCACAGTGTGAATGCCCT 910
Db 281 PheSerTrpLeuArgTrpAsnSerTrpLeuGlnGlnSerLeuHisArgSerGlnCysPro 300
QY 911 TCCGAGCGGTATATCTCTTCAGTGTTCCTCCAGTGCAGAGGCGGCGGCGGCGGCGG 970
Db 301 SerGlnArgTyrLeuSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATGCTGGAGGGGGCGGCTCGGATGAGGAGGCGCTTGGCAAGTGAAGTCTGACATTC 1030
Db 321 IleValGlyValAlaLeuAlaSerHisSerTrpTrpTrpGlnValSerLeuHisPhe 340
QY 1031 GGCACCAACCAATCTGTGAGGAGCAGCTTATGACCCAGGAGGAGTGTCTACTCCGCC 1090
Db 341 GlyThrThrHisIleCysGlyGlyThrLeuLeuAspHisGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTGTGACCCGGAGAGAGGTCCTGGAGGCTGGAAAGTGTACCGGCGCAC 1150
Db 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACCTGACCAAGTTCGAGGAGGAGGCTTCATTCGCGAGATCATCATCAAGCAAT 1210
Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerLeuHisGlnGlnLeuSerHisSer 400
QY 1211 TACACGATGAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1270
Db 401 TyrThrAspGlnGlnAspAspTrpAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY 1271 CTGTCCGCTCAATCAACCTGCTGCTCCCTCCCATGATGAGAGAGAGAGAGAGAGAG 1330
Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGCGGATCAAGAGGCTTGGCAAGAGAGAGAGAGAGAGAGAGAGAGATCCGCC 1390
Db 441 GlnThrCysTrpIleHisGlyPheGlyLysThrArgGlnThrAspAspLysTrpHisSerPro 460
QY 1391 TTCTCCGAGAGGAG 1450
Db 461 PheLeuArgGlnValGlnValAlaLeuLeuLeuAspPheLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACAGTTACTTACCCCAAGAGATGTGTGCTGGGAGCTTCTGCGGGGCGAGAGAC 1510
Db 481 TyrAspSerTrpLeuThrTrpArgMetCysAlaGlyAspLeuValArgGlyArgAsp 500
QY 1511 TCCGCGAG 1570
Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnAspTrpTrpLeu 520
QY 1571 GAGGTGTACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACAC 1630
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY 1631 AAAGTGAACAGAGTTCTTCTGATTTTACAGCAAGATGAGAGAGAGAGAGAGAGAGAG 1690
Db 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlnSerGlnValArgPheIle 560

QY 1691 AATTC 1696
|||||
Db 561 LySer 562

RESULT 4

AD110393
ID AD110393 standard; protein; 562 AA.

XX AD110393;

DT 22-APR-2004 (first entry)

XX Human cell surface protease #12.

XX therapeutic agent; plasmin; protease specific antigen; PSA;

KW cell-surface protease-associated disease; cancer; ocular disease;

KW cardiovascular disease; chronic inflammatory disease; wound;

KW circulatory disorder; dermatological disorder; rheumatoid arthritis;

KW psoriasis; diabetic retinopathy; pterygium;

KW excimer laser surgery scarring; glaucoma filtering surgery scarring;

KW macular degeneration; cress syndrome; solid neoplasm; vascular tumour;

KW melanoma; Kaposi's sarcoma; human; cell surface protease.

XX Homo sapiens.

XX MO20295007-A2.

XX 28-NOV-2002.

XX 23-MAY-2002; 2002MO-US016819.

XX 23-MAY-2001; 2001US-0293267P.

XX (CORV-) CORVAS INT INC.

PI Madison BL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV;

XX WPI; 2003-221280/21.

DR N-PSDB; AD110392.

PT Novel conjugate useful for treating cell-surface protease-associated

XX substrate linked to it optionally by a peptidic linker.

XX Claim 9; SEQ ID NO 24; 581bp; English.

CC The invention comprises a conjugate that consists of a therapeutic agent

CC and a peptide substrate (optionally linked via linker). The peptide

CC substrate is proteolytically cleaved by a cell surface protease pr a

CC soluble, released or shed form of it, to liberate the therapeutic agent,

CC the conjugate of the invention is not substantially cleaved by plasmin or

CC protease specific antigen (PSA). The conjugate of the invention is useful

CC for treating a cell-surface protease-associated disease such as: cancer,

CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,

CC wounds, circulatory disorders, dermatological disorders, rheumatoid

CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,

CC scarring from excimer laser surgery, scarring from glaucoma filtering

CC surgery, macular degeneration, cress syndromes, solid neoplasms, vascular

CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence

XX represents a human cell surface protease.

XX SQ Sequence 562 AA;

XX Alignment Scores:

Score: 2.77e-180 Length: 562

Percent Similarity: 99.84 Matches: 561

Best Local Similarity: 99.84 Mismatches: 0

Query Match: 89.74 Gaps: 1

DB: 7 Indels: 0

US-10-806-370-11 (1-1748) X AD110393 (1-562)

QY 11 ATGAGAGAGGACAGCCAGGGAATGATCTTCAGCAAGAAACCTTGAGTGCATCT 70

Db 1 MetGluArgAspSerHisIleGlyAsenHisSerProHisArgThrProSerHisIleGlyAlaSer 20

QY 71 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCAGGCGGGGATCTCCAGCCCAAGGATCT 130

Db 21 ProHisArgHisSerProHisArgThrProProHisArgHisSerProHisArgHisSer 40

QY 131 CCAGCCCAAGGATCTCCAGCTGGGACACCTCCGAGCGGGGATCTCCAGCCCAAGGATCT 190

Db 41 ProHisArgHisSerProHisArgThrProProHisArgHisSerProHisArgHisSer 60

QY 191 CCAGCTGTACACCTCCAGGCGGGGATCTCCAGGCGGGGATCTCCAGCCCAAGGATCT 250

Db 61 ProHisArgHisSerProProHisArgThrProProHisArgHisSerProHisArgHisSer 80

QY 251 CCAGCCCGGGGATCTCCGAGCTGGGATCTCCAGCTGGGATCTCCAGCTGGGATCTCA 310

Db 81 ProHisArgHisSerProHisArgThrProProHisArgHisSerProHisArgHisSer 100

QY 311 TCGCCAGGTCAGGCTCGGTGACCAACCTCCCAAGGATCTGTGTAGAGCAACA 370

Db 101 SerHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 120

QY 371 CCAGTGGGGCTGTACCATCTCCATCTCTCCAGGTCAGGTCAGGTCAGGTCAGGTC 430

Db 121 ProHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 140

QY 431 ACCAGGAGAGCCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 490

Db 141 ThrArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 160

QY 491 CTACCGCTCATCGGTCGTCCT 550

Db 161 LeuProHisSerProHisArgThrProHisArgThrProHisArgHisSer 180

QY 551 TTCAGGTCGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 610

Db 181 PheHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 200

QY 611 AAGCAGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 670

Db 201 LysHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 220

QY 671 TGCCTGAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 730

Db 221 CysHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 240

QY 731 TGCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 790

Db 241 TrpLeuProHisSerProHisArgThrProHisArgThrProHisArgHisSer 260

QY 791 CTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 850

Db 261 LeuHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 280

QY 851 TTTCATCTTGAATCAACTCCAGCCTCCAGGAAAGCTCCAGGTCGTCGTCGTCGTCGTC 910

Db 281 PheSerHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 300

QY 911 TCCAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 970

Db 301 SerHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 320

QY 971 ATGCTGGAGGAGGCTCGGCTCGGATGAGGAGGCTGGCAAGGAGGTCGTCGTCGTCGTC 1030

Db 321 IleValHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 340

QY 1031 GGCACACCAACATCTGTGAGGACCGCTGATGACCCAGTGGTGTCTCACTGCGGCC 1090

Db 341 GlyThrHisArgHisSerProHisArgThrProHisArgThrProHisArgHisSer 360

QY 1091 CACTGCTTCTTCTGACCCGAGGAGAGGTCCTGAGGAGGTCGAAAGTGTACCGCGGACCC 1150


```

Db      361 HiscyshpheValThrArgGluValIleuGluGlyTrpLysValTyrAlaGlyThr 380
QY      1151 AGCAACCTGCACCGCTTGCTGAGGAGCGCTCCATTCGCGAGATCATCATCAAGCAAT 1210
Db      381 SerAsnLeuHisGlnLeuProGluAlaIleAsnLeuIleIleIleAsnSerAsn 400
QY      1211 TACACCGATGAGGAGGAGCACTATGACATGCGCCCTCATGGGGCTGTCCAGCCCTGACC 1270
Db      401 TyrThrAspGluGluAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY      1271 CTGTCCGCTGCATCCACCCCTGCTTGCTGCTCCCATGATGAGCAGACCTTTAGCCTCAAT 1310
Db      421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY      1331 GAGACCTGCTGATCCACAGCGCTTTGGCAAGACGAGGAGACAGATGACAGACATCCCTCC 1390
Db      441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro 460
QY      1391 TTCTCTCCGGGAGGTGAGGTCAATCTCATGCACTTCAAGAAATGCATGACTATTGCTC 1450
Db      461 PheLeuArgGluValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal 480
QY      1451 TATACAGATTACCTTACCCCAAGATGATGTGTCTGGGAGCTTCGTTGGGGGAGAGAC 1510
Db      481 TyrAspTyrLeuThrProArgMetCysAlaGlyAspLeuArgGlyLysArgAsp 500
QY      1511 TCCTGCAGGAGAGACAGCGGGGGGCTTGTGTGTGAGCAGAAACAACCGTGTACTCTG 1570
Db      501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY      1571 GCAGGTGCACACAGCTGGGGGACAGGCTGTGTGCGCAGAGAAACAACCTGGTGTACACC 1630
Db      521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY      1631 AAAGTGCAGAGAGTTCTTCCTGATTTTACAGCAAGATGAGAGCGAGGTSCATTGAGA 1690
Db      541 LysValThrGluValLeuProTrpIleTyrSerLysMetGlnSerGluValArgPheIle 560
QY      1691 AATATC 1696
Db      561 LysSer 562

```

```

PR      23-MAY-2002; 2002US-00156214.
XX
PA      (CORV-) CORVAS INT INC.
XX
PI      Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Stew DV;
XX
DR      MPI; 2004-190126/18.
XX
DR      N-PSDB; ADJ46917.
XX
PT      Conjugate useful for treating e.g. cancer, cell-surface protease-
PT      associated diseases, comprising a peptidic substrate or nucleic acid
PT      substrate linked to a therapeutic agent through a linker.
XX
PS      Claim 9; SEQ ID NO 24; 361pp; English.
XX
CC      The invention relates to a conjugate comprising a therapeutic agent and a
CC      peptidic substrate or nucleic acid substrate linked to the agent
CC      optionally through a linker or peptidic linker, where the peptidic
CC      substrate is proteolytically cleaved by a cell surface protease or a
CC      soluble, released or shed form conjugate to liberate the agent and the
CC      conjugate is not substantially cleaved by plasmin or prostate specific
CC      antigen (PSA). The conjugate is useful for treating a disease, which
CC      involves administering a conjugate to a subject, where the disease is
CC      preferably a proliferative diseases or a cell-surface protease-associated
CC      disease. The diseases include autoimmune diseases, inflammatory diseases,
CC      infectious diseases and endocrine diseases. The conjugate is useful for
CC      treating a cell-surface protease-associated disease, which involves
CC      administering a conjugate comprising an agent and a peptidic substrate to
CC      a subject exhibiting symptoms of a cell-surface protease-associated
CC      disorder, where the disease is selected from cancer, ocular disorders,
CC      cardiovascular disorders, chronic inflammatory disorders, restenosis, wounds,
CC      circulatory disorders, dermatological disorders, rheumatoid
CC      arthritis, psoriasis, diabetic retinopathies, scarring from laser
CC      surgery, scarring from glaucoma filtering surgery, macular degeneration,
CC      CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
CC      vascular tumours such as lung cancer, colon cancer, prostate cancer,
CC      melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
CC      serine protease (MSP)-related polypeptide of the invention.
XX
SQ      Sequence 562 AA:

```

Alignment Scores:

	Pred. No.:	Length:	Matches:
Score:	2,77e-180	562	561
Percent Similarity:	2991.00		
Best Local Similarity:	99.8%	Conservative:	0
Query Match:	99.8%	Mismatches:	1
DB:	89.7%	Indels:	0
		Gaps:	0

```

US-10-806-370-11 (1-1748) x ADJ46917 (1-562)
QY      11 ATGAGAGGAGCAGCGCGGAATGCATCTCCAGCAAGAACACCTTGAGCATCT 70
Db      1 MeGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY      71 CCAGCCGAGCATCTCCAGCTGGGACACCTCCAGCCGGGCGCATCTCCAGCCGAGCATCT 130
Db      21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY      131 CCAGCCGAGCATCTCCAGCTGGGACACCTCCGGGCGGGGCGCATCTCCAGCCGAGCATCT 190
Db      41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY      191 CCAGCTGTACACCTCCAGGCGGGGACATCTCCAGCGGGGCGCATCTCCAGCCGAGCATCT 250
Db      61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY      251 CCAGCCGGGCGCATCTCCGCTGGGACATCTTCCAGGTCTCATCTCGGAGGTCATCA 310
Db      81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
QY      311 TCCGCGAGGTGAGCTGTGAGCAACCTCCCAACGAGATGATCTGTAGAGCAACA 370

```

Db	101	SerialaAgSerAlaSerValThrThrsSerProThrArgValTyrLeuValAlaArgAlaThr	120
OY	371	CCAGGCGGGGGCTGTACCCCATCCGATCATCTCCCTGCGAGGTACAGCACCAAGGCGC	430
Db	121	ProValGlyAlaValAlaProIleArgSerSerProAlaAgSerAlaProAlaThrArgAla	140
OY	431	ACCAAGGAGACCCCAAGGTACAGAGCTGGCCCAAGTTCACTCGCGGGAAGGCCAAGACAG	490
Db	141	ThraArgGluSerProGlyThrSerLeuProLysPheThrTrpArgGluGlyLysGln	160
OY	491	CTACCGCATCATCGGGTGGCTGCTCTCCATATGCGCCGTGGTGGTTTCCCTCATCATCTGC	550
Db	161	LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleLeu	180
OY	551	TTCCAGTTCTGACAGAGGCCACACAGAGATCAAGATCAAGAGACAGAGGAGAGAGCTGGCC	610
Db	181	PheGlnPheTrpGlnGlyIleThrGlyIleArgTyrLysGlnArgGluSerCysPro	200
OY	611	AAGCAGCGCTGTTCCGTGACCGGGGTGGTGAATGCAAGCTGAAGAAGTGAAGAGCTGGGC	670
Db	201	LysAlaValAlaArgCysAspGlyValAlaAspCysLysLeuLysSerAspGlyLeuGly	220
OY	671	TGGGAGAGATTGACGTGGGACAAAGCTCGCTTAATCTACTCGGGGCGCCCATAG	730
Db	221	CysValArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHisGln	240
OY	731	TGGCTTCCATCTGTAGACCAACAATCGAATGACTCTACTCAGAGAAGACCTGCCAGAG	790
Db	241	TrpLeuProIleCysSerSerAlaThrAspSerTyrSerGluLysThrCysGlnGln	260
OY	791	CTGGATTTCAGAGAGCTCTCACCGGACAAACCGAGGTTGCCACAGGGATTTTGGCAACAGC	850
Db	261	LeuGlyPheGluSerAlaHisArgThrArgValAlaHisArgAspPheAlaHisSer	280
OY	851	TTCTCAATCTTGAAGATACAACTCCACCAATCCAGAAAGCGTCCACAGGCTGAAAGCCCT	910
Db	281	PheSerIleLeuAlaGlyTrpAspSerThrIleGlnGluSerLeuHisArgSerGlyCysPro	300
OY	911	TCCACCGGGATATATCTCCCTCCAGTGTCCCACTGCGAGCTAGAGGCGCATGACCGAGGCG	970
Db	301	SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg	320
OY	971	ATCGTGGAGAGGGCGCCTGGCCTCGGATACGAAGTGGCCTTGGCAAGTGAATCTGCACTTC	1030
Db	321	IleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe	340
OY	1031	GGACACACCCATCATGTGGAGAGGACGGCTATTCAGGCGGAGTGGGTGTCACTGCGCGC	1090
Db	341	GlyThrThrHisIleLeuCysGlyValGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla	360
OY	1091	CAGTGGCTTCTTGAGACCCGGAGAGAAAGTCTGTGAGGCGCTGGAAAGGTGACCGGCGAC	1150
Db	361	HisCysPhePheValThrArgLysValValLeuGluGlyTrpLysValTyrAlaIleLysThr	380
OY	1151	AGCAACCTGCACCAAGTTGCTGAGGACGCTTCATTCGCGAGATATATCAACAGCAAT	1210
Db	381	SerAlaLeuHisGlnLeuProGlyAlaIleSerIleAlaGluIleIleIleAlaSerAsn	400
OY	1211	TACACCGATGAGAGAGAGACATATACATCGCCCTCAACGCGCTGTCCAAAGCCCGTAC	1270
Db	401	TyrThrAspGluGlnAspAspArgLysAspIleAlaAlaMetAspArgLeuSerLysProLeuThr	420
OY	1271	CTGTCCGCTCATCCACCCCTGCTGCTGCCCTCCCAAGATGACAGACCTTTAGCTCATAT	1330
Db	421	LeuSerAlaHisIleIleHisProIleCysLeuProMetHisGlyGlnThrPheSerLeuAsn	440
OY	1331	GAGACCTGCTGGATGACAGGCTTTTGACAAAGACAGGAGACAGATGACAAAGATCCCC	1390
Db	441	GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysThrSerPro	460
OY	1391	TTCCCTCCCGGAGGCTCAGGTCAATCTCATCGACTTCAAGAAATGCAAGATCACTTGGATC	1450
Db	461	PheLeuArgGlyValAlaGlnValAlaSerLeuIleAspPheLysCysAspAspArgLysVal	480

QY	1451	TATGACAGTTA	CCTTACCCCAAGATGATGTGTGCTGGGGA	CCTTGTGGGGGCAGAGAC	1510
Db	481	TyzAsperTyrrLeuTrpProArgMetMetCysAlaIlyAspLeuArgGlyArgAsp			500
QY	1511	TCCTGACAGGAGAGACAGCGGGGGGGCCCTTGTGTGTGAGACAGAA	CAACCGCTGGTACTCTG		1570
Db	501	SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnAlaGlyTrpLeu			520
QY	1571	GCAGGTCTCACACAGCTGGGGGCACAGGCTGTGGCCAGAAACAA	ACCTGTGTGTACACC		1630
Db	521	AlaGlyAlaThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr			540
QY	1631	AAATGACAGAGAGCTTCTCCCTGGATTATACAGCAAGATGGAGAGCCAGGTCGCA	TTTCAGA		1690
Db	541	LysValThrGluValLeuProTrpIleTyrSerLysMetGlnSerGluValArgPheIle			560
QY	1691	AAATCC	1696		
Db	561	LysSer	562		
RESULT 6					
ADY50146					
ID	ADY50146	standard; protein; 562 AA.			
XX	AC	ADY50146;			
XX	DT	19-MAY-2005 (first entry)			
XX	DE	Human endotheliase-2L.			
XX	KW	Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;			
KW	KW	Ophthalmological; Cardiovascular-Gen; Vuninary; Antiinflammatory;			
KW	KW	Vasectropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;			
KW	KW	rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;			
XX	XX	dermatological disease; cancer; neoplasm; endotheliase-2; enzyme.			
OS	OS	Homo sapiens.			
XX	XX	WO2005019270-A2.			
XX	XX	03-MAR-2005.			
XX	XX	12-AUG-2004; 2004WO-US026148.			
XX	XX	14-AUG-2003; 2003US-0495005P.			
PR	PR	14-NOV-2003; 2003US-0520164P.			
XX	XX	(DYAX-) DYAX CORP.			
PA	PA	(DEND-) DENDREON CORP.			
XX	XX	Madison EL, Nixon A;			
XX	XX	WPI; 2005-202609/21.			
DR	DR	N-PSDB; ADY50145.			
PT	PT	Novel proteain capable of inhibiting endotheliase-2, useful for treating			
XX	XX	or preventing angiogenesis related disorder e.g. cancer.			
PS	PS	Disclosure; SEQ ID NO 94; 157bp; English.			
XX	XX	The invention relates to an isolated protein (I) which comprises a heavy			
CC	CC	chain (HC) immunoglobulin variable domain sequence and a light chain (LC)			
CC	CC	immunoglobulin variable domain sequence, where the first and second			
CC	CC	immunoglobulin variable domain sequences from an antigen binding site			
CC	CC	that specifically binds to human endotheliase-2. (I) is useful for			
CC	CC	detecting an endotheliase or endotheliase activity in a sample, for			
CC	CC	modulating an activity of an ET2-expressing cell, for modulating			
CC	CC	proteolysis, for killing or inhibiting growth of a cell, for detecting			
CC	CC	endotheliase in a subject, for modulating endotheliase activity in a			
CC	CC	subject, for treating or preventing a disorder characterized by unwanted			
CC	CC	angiogenesis in a subject. The disorder is chosen from rheumatoid			
CC	CC	arthritis, psoriasis, diabetic retinopathies, ocular disorder such as			

CC pterygii recurrence, scarring excimer laser surgery and glaucoma
CC filtering surgery, cardiovascular disorders, chronic inflammatory
CC disorders, wound repair, circulatory disorders, crest syndromes,
CC dermatological disorders and cancers. The present sequence represents the
CC human endothelialase-25.

XX Sequence 562 AA;

Alignment Scores:

Pred. No.: 2.77e-180 Length: 562
Score: 2991.00 Matches: 561
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 89.7% Indels: 0
DB: 9 Gaps: 0

US-10-806-370-11 (1-1748) x ADY50146 (1-562)

QY 11 ATGGAGAGGAGCAGACGAGGAAATGCAATCTCCAGCAAGAACCTTCAGCTGAGCATCT 70
Db 1 MetGluArgAspSerHisLeuValAsnAlaSerProAlaGlnProSerAlaGlyAlaSer 20
QY 71 CCAAGCCAGGATCTCCAGCTGGAGACACCTCCAGCCGGGATCTCCAGCCAGGATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAAGCCAGGATCTCCAGCTGGAGACACCTCCAGCCGGGATCTCCAGCCAGGATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAAGCTGTACATCCAGGCGGGGATCTCCAGCCGGGATCTCCAGCCAGGATCT 250
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAAGCCGGGATCTCCAGCTGGAGACATCTTCAGGCTCTCATCCGAGGATCTCA 310
Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerHisSerSerGlyArgSerSer 100
QY 311 TCCGCCAGGTCAGCTCGGTGACACCTCCCAACAGAGTGAACCTTTAGAGACA 370
Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAAGTGGGGGTGAACCATCCAGATCATCTCCAGGTCAGAGCCAGCAACAGGACC 430
Db 121 ProAlaGlyAlaValProAlaArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
QY 431 ACCAGGAGAGCCAGATGACAGGCTCCAGTTCACCTGCGGAGAGGCGCAGAAAGCAG 490
Db 141 ThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlyGlnLysGln 160
QY 491 CTACCGGTCATCGGAGTGGCTCTCTCTCATTTGCTGTTGGTTCCTCATCATCTCTC 550
Db 161 LeuProLeuLysGlyCysValLeuLeuLeuLeuValAlaSerLeuLeuLeuLeu 180
QY 551 TTCCAGTTCGAGGAGGCGGACAGAGGATGACGTAACAAGAGGAGGAGGAGGAGGAGG 610
Db 181 PheGlnPheTrpGlnGlyHisThrGlyTyrLeuGlyLysGlnGlnArgGlnSerCysPro 200
QY 611 AAGCAGCGTCTTCGTCGAGCGGAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 670
Db 201 LysHisAlaValAlaGlyCysAspGlyValValAspCysLysLeuLysSerAspGlnLeuGly 220
QY 671 TGGCTGAGGTTTGAAGTGAAGTCTCTGCTTAAATCTAATCTGAGTCTCTCCATCAG 730
Db 221 CysValAlaArgPheAspTrpAspLysSerLeuLysLysLysLysLysLysLysLysLys 240
QY 731 TGGCTTCCATCTGAGAGCAAGTGAAGTCTCTAAGAGAGAGAGAGAGAGAGAGAGAGAG 790
Db 241 TrpLeuProLysCysSerSerAsnTrpAsnAspSerLysSerGlnLysLysLysLysLys 260
QY 791 CTGGGTTTGAAGTGTCTACCGAGCAACGAGGTTGGCCAGAGGATTTTCCAAACAGC 850
Db 261 LeuGlyPheGlnSerAlaHisAlaGlyThrGlnValAlaHisValGlyAspPheAlaAspSer 280

QY 851 TTCTCAATCTTGAGATACACTCCACCATCCAGGAAGCCTCCACAGCTGAATGCGCT 910
Db 281 PheSerLeuLeuArgLysAsnSerThrLysGlnLysSerLeuHisAspSerGlnCysPro 300
QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCACTGCGAGTGAAGGAGCCATGACCGGCGG 970
Db 301 SerGlnArgTyrLysSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGGGGCGCTGCGCTCGGATGACAAAGTGGCTTGGCAAGTGAAGTTCACATTC 1030
Db 321 IleValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe 340
QY 1031 GGCACACCCACATCTGTGAGGAGCGCTCATGAGCCCGAGTGGGTGCTCACTGCGCGCC 1090
Db 341 GlyThrThrHisLysCysGlyGlyThrLeuLeuHisAspAlaGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTGTGACCCGAGAAAGTCTCGAGGAGGCTGGAAGTGTACCGCGGACCC 1150
Db 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACTGCAACAGTTGCTGAGGACAGCTCCATTCGCCAGATCATCATCAACAGCAAT 1210
Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleHisSerAsn 400
QY 1211 TACACCGATGAGAGGACGACTATGACATGCGCTTCATGGGCGTTCGCAAGCCCTGACC 1270
Db 401 TyrThrAspGlnGlnAspAspTrpAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY 1271 CTGTCGCTCATCATCCAGCTGCTGCTGCTCCCATGTCATGGAACAGACTTTAGCTCAT 1330
Db 421 LeuSerAlaHisLysIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGCTGATCAGAGGCTTTGGCAAGACGAGGAGACAGATGACAAAGATCTCCCC 1390
Db 441 GlnThrCysTrpLysThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
QY 1391 TTCTCTCGGAGAGTGGAGTCAATCTCATGACTTCAAGAAATGCAATGACTTACTTGTC 1450
Db 461 PheLeuArgGlnValGlnValAsnLeuLeuAspPheLysLysCysAsnAspTyrLeuVal 480
QY 1451 TATACAGTATACCTTACCCCAAGGATGATGTGCTGGGACCTTCGTGGGGGAGAGAC 1510
Db 481 TyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeuAlaGlyGlyArgAsp 500
QY 1511 TCCTGCCAGGAGACAGCGGGGCTCTTGTCTGTGAGCAAGAAACAACGCTGTATCTG 1570
Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY 1571 GCAGGTGTACACAGCTGGGCGACAGGCTGTGGCCAGAGAAACAACCTGTGTGTACACC 1630
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY 1631 AAAGTGCAGAAAGTCTTCCCTCGATTTACAGCAAGATGAGAGCGAGTGCATTCAGA 1690
Db 541 LysValThrGlnValLeuProTrpLysSerLysMetGlnSerGlnValArgPheIle 560
QY 1691 AAATCC 1696
Db 561 LysSer 562

RESULT 7
AAE01944 ID AAE01944 standard; protein; 688 AA.
XX AAE01944;
AC AAE01944;
DE 31-JUL-2001 (first entry)
XX Human transmembrane serine protease (Endothelialase 2-L) protein.
DE Human; endothelialase 2-L; protease domain; cytosolic; vlnitary; wound;
KW neotrophic; periodontitis; dermatological disorder; gene therapy; scar;

KM Angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
KM chronic inflammatory disease; ocular disorder; circulatory disorder;
KM celiac syndrome; atherosclerosis; haemangiomas; diabetes mellitus;
KM liver cirrhosis, osteoradionecrosis, systemic sclerosis; oesophageal;
KM inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
KM systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
KM transmembrane serine protease.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 321..688
FT /label= Protease_domain
XX MO200136604-A2.
XX 25-MAY-2001.
XX 17-NOV-2000; 2000MO-US011803.
XX 18-NOV-1999; 99US-0166391P.
XX 22-SEP-2000; 2000US-0234840P.
XX (CORV-) CORVAS INT INC.
XX Madison BL, Ong EO;
XX MPI; 2001-336001/35.
XX N-PSDB; AAD05797.
PT New nucleic acid encoding a protein comprising endothelinase activity
PT useful in the prevention and treatment of e.g. vascular malformations,
PT cardiovascular disorders, and chronic inflammatory disease.
XX Claim 4; Page 142-143; 152pp; English.
XX The present sequence is human long form transmembrane serine protease
XX (Endothelinase 2-) protein. Endothelinase 2 is a type-II membrane-type
XX serine protease which has a transmembrane domain at the N-terminus,
XX followed by a single low density lipoprotein-A receptor domain and a
XX single scavenger-receptor cysteine-rich domain. The C-terminus of the
XX endothelinase 2 contains the trypsin-like serine protease domain
XX characterised by the catalytic triad residues in 3 highly conserved
XX regions of the catalytic domain. In addition 3 repetitive sequence are
XX found just before the transmembrane domain and represent a sequence motif
XX for N-myristoylation modification. The invention relates to an
XX endothelinase protein, endothelinase protease domain and their
XX corresponding nucleic acid molecules. An endothelinase protein or protease
XX domain of it is useful for the treatment and diagnosis of disorders
XX associated with aberrant angiogenesis or undesired neovascularisation.
XX The undesired angiogenesis is associated with disorders selected from
XX solid neoplasms, vascular malformations and cardiovascular disorders such
XX as angiofibroma, angiolipoma, atherosclerosis, restenosis/reperfusion
XX injury, arteriovenous malformations, haemangiomas and vascular
XX adhesions, dyschondroplasia with vascular hamartomas (Patau's
XX syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
XX syndrome) and Von Hippel Lindau syndrome, chronic inflammatory diseases
XX such as diabetes mellitus, haemophilic joints, inflammatory bowel
XX disease, nonhealing fractures, periodontitis, psoriasis, rheumatoid
XX arthritis, venous stasis ulcers, granuloma-in-burns, hypertrophic scars,
XX liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
XX granuloma and systemic sclerosis and aberrant wound repairs, circulatory
XX disorders Raynaud's phenomenon, crest syndromes such as calcinosis,
XX oesophageal, dysmotility, sclerodactyly and telegiectasis, dermatological
XX disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
XX vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
XX stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
XX and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
XX caused by ocular neovascular disease, corneal graft neovascularisation,
XX macular degeneration, retinopathy of prematurity, retrolental fibroplasia
XX and corneal neovascularisation. The nucleic acids of the invention are
XX also used in gene therapy. The invention also provides method for
XX screening compounds that modulate angiogenesis

XX SQ Sequence 688 AA;
XX Alignment Scores:
XX Pred. No.: 1,67e-178 Length: 688
XX Score: 2963.00 Matches: 554
XX Percent Similarity: 99.6% Conservative: 2
XX Best Local Similarity: 99.3% Mismatches: 2
XX Query Match: 88.9% Indels: 0
XX DB: 4 Gaps: 0
US-10-806-370-11 (1-1748) x AAE01944 (1-688)
QY 11 ATGGAGAGAGGACAGCCAGGAAATGCACTCCAGCAAGAAACACCTTCAGCTGAGCATCT 70
Db 1 MetGluArgAspSerHisGLYAsnAlaSerProAlaArgHisProSerAlaGLYAlaSer 20
QY 71 CCAGCCCAAGCATCTCCAGCTGGGACACCTTCAGGCCGGGCATCTCCAGCCCAAGCATCT 130
Db 21 ProAlaGlnAlaSerProAlaGLYThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCCAAGCATCTCCAGCTGGGACACCTTCAGGCCGGGCATCTCCAGCCCAAGCATCT 190
Db 41 ProAlaGlnAlaSerProAlaGLYThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGTACACCTCCAGGCCGGGCATCTCCAGGCCGGGCATCTCCAGCCCAAGCATCT 250
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAGCCCGGGCATCTCCAGCTGGGACATCTTCAGGTCCTTCATCCGGCAGGTATCA 310
Db 81 ProAlaArgAlaSerProAlaAlaSerAlaSerAlaSerAlaSerAlaSerAlaSer 100
QY 311 TCGCCAGGTCAGGCTGAGTGAACAACCTCCCAACAGAGTACTTGTAGAGCAACA 370
Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValThrLeuValArgAlaThr 120
QY 371 CCAAGTGGGGGCTGTACCCATCCATCCATCTCTCCAGTCCAGGACCAAGCAAGGAGCC 430
Db 121 ProValGlyAlaValProLeuAlaArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
QY 431 ACCAGGAGAGCCCAAGGTACGAGCTCCCAAGTTCACTTCGGCGGAGGCGCAGAACAG 490
Db 141 ThrArgGluSerProGlyThrThrSerLeuProLysPheThrArgGluGlnLysGln 160
QY 491 CTACCGGTCAATCGGGGCGGTCTCTCTCTCATGGCCCTGGGGTTCGCTCATCATCTCT 550
Db 161 LeuProLeuLleGlyCysValLeuLeuLeuLleAlaLeuValAlaSerLeuLleLeu 180
QY 551 TTCAGTTCGTGGCAGGCGCACACAGGATCAGGTACAGAGCAGAGAGAGTGTCC 610
Db 181 PheLlnHisThrPrlGlnGlnLysAlaThrGlyLleArgGlyLysGlnAlaGlyLysCysPro 200
QY 611 AAGACGCTGTTCGCTGTGACCGGGGTGTGACCTGCAGCTGAAAGATGACAGCTGGC 670
Db 201 LysHisAlaValArgCysAspArgValValAspCysLysLeuLysSerAspArgLleGly 220
QY 671 TGGGTGAGGTTTGTGCTGGGACAGTCTCTGTTAAATTAATTAATTAATTAATTAATTA 730
Db 221 CysValArgPheAspThrPheLysSerLeuLeuLysAlaLysSerGlySerHisGln 240
QY 731 TGGCTTCGATCTGTGACAGCAACCTGGAATGACTCTTACTCAGAGAAAGCTGGCAGAG 790
Db 241 TrpLeuProLleCysSerSerSerAlaThrPheAspSerGlySerGlyGln 260
QY 791 CTGGGTTTCAGAGAGTGTCCACCGGACCAAGGATTCCTCCACAGGATTTTCCCAACAG 850
Db 261 LeuGlyPheGluSerAlaHisAlaArgThrThrGluValAlaHisArgAspPheAlaAsnSer 280
QY 851 TTTCATCTTGTAGATACACTCCACCATCCAGGAAGAGCTCCACAGAGTGAATGCCCT 910
Db 281 PheSerLleLeuArgLysAsnSerThrLleGlnLysSerLeuHisAlaArgSerGluCysPro 300

QY 911 TCCAGCGGATATATCTCCCTCCAGTGTTCCTCCAGTGGAGGCGCATGACCGGCGG 970
 DB 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuAlaGlyAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGGGGCGCTGCGCTCGATAGCAATGGCGCTTGGCAAGTGAAGTCTGCACCTC 1030
 DB 321 lIeValGlyGlyAlaLeuAlaSerAspSerIleTyrProTglnValSerLeuHisPhe 340
 QY 1031 GGCACCAACCCATCTGTGGAGGCAAGCTCATTCACCGCCGATGGGTCTCTACCTCCGCC 1090
 DB 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTyrValLeuThrAlaAla 360
 QY 1091 CACTGCTTCTCTGAGACCGGGAGAGAGTCTGGAGGGCTGGAAGGTGTAACCGGCGCAC 1150
 DB 361 HisCysPhePheValThrArgGlnValLeuGlnGlyTyrIleValIleGlyThr 380
 QY 1151 AGCAACCTGCACACGTTGCTGAGGACGCTTCATTCGATCGGATCATCATCAAGCAAT 1210
 DB 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
 QY 1211 TACACCGATGAGAGAGCACTATGACATCGCCCTCATCGGCTGTCCAGCCCTGACC 1270
 DB 401 TyrThrAspGlnGluAspAspTyrAspIleAlaLeuMetArgLeuSerIleProLeuThr 420
 QY 1271 CTGTCGGCTGACATCCACCCCTGCTGCTCCCATGATGACAGACCTTTAGCTCAAT 1330
 DB 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGCTGATATCAGAGCTTTGGCAAGACAGGAGAGACAGATGACATGCCCC 1390
 DB 441 GluThrCysTyrIleThrGlyPheGlyIleThrArgGlnTyrAspAspIleThrSerPro 460
 QY 1391 TTCTCTCGGGAGGTGACAGTCAATCTCATTCGACTTCAGAAATGATGACTTGTGTC 1450
 DB 461 PheLeuArgGlyValGlnValAsnLeuIleAspPheIleValCysAsnAspTyrLeuVal 480
 QY 1451 TATGACATTTACCTTACCCCAAGAGTATGATGTGTGGGACCTTCGGGGGAGAGAC 1510
 DB 481 TyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeuAlaGlyGlyIleArgAsp 500
 QY 1511 TCCTGCGGAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAAACCGCTGTGACCTG 1570
 DB 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTyrIleu 520
 QY 1571 GCAGGTGTACCAAGCTGGGGCAAGGCTGTGGCCAGAGAAACAACCTGTGTGTACCC 1630
 DB 521 AlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnIleProGlyValIleThr 540
 QY 1631 AAAGTGAAGAGTCTTCCCTGATTTTACGCAAGATGAGAGCGAGGTGCGA 1684
 DB 541 LysValThrGlnValLeuProTyrIleTyrSerIleMetGluAsnArgAlaGln 558
 RESULT 8
 AD110395 ID AD110395 standard; protein: 688 AA.
 XX AD110395;
 AC
 DT 22-APR-2004 (first entry)
 XX
 DE Human cell surface protease #13.
 XX
 KW therapeutic agent; plasmin; protease specific antigen; PSA;
 KW cell-surface protease-associated disease; cancer; ocular disease;
 KW cardiovascular disease; chronic inflammatory disease; wound;
 KW circulatory disorder; dermatological disorder; rheumatoid arthritis;
 KW psoriasis; diabetic retinopathy; pterygium;
 KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; cret syndrome; solid neoplasm; vascular tumour;
 KW melanoma; Kaposi's sarcoma; human; cell surface protease.
 XX
 OS Homo sapiens.

PN W0200295007-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 23-MAY-2002; 2002WO-US016819.
 XX
 PR 23-MAY-2001; 2001US-0293267P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Madison EL, Sempke JE, Vlaux GP, Kemp SJ, Komandla M, Slev DV;
 XX
 DR WPI: 2003-221280/21.
 DR N-PSDB; AD110394.
 PT Novel conjugate useful for treating cell-surface protease-associated
 PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
 PT substrate linked to it optionally by a peptidic linker.
 XX
 PS Claim 9; SEQ ID NO 26; 581bp; English.
 XX
 CC The invention comprises a conjugate that consists of a therapeutic agent
 CC and a peptide substrate (optionally linked via linker). The peptide
 CC substrate is proteolytically cleaved by a cell surface protease pr a
 CC soluble, released or shed form of it, to liberate the therapeutic agent,
 CC the conjugate of the invention is not substantially cleaved by plasmin or
 CC protease specific antigen (PSA). The conjugate of the invention is useful
 CC for treating a cell-surface protease-associated disease such as: cancer;
 CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
 CC wound, circulatory disorders, dermatological disorders, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
 CC scarring from excimer laser surgery, scarring from glaucoma filtering
 CC surgery, macular degeneration, cret syndromes, solid neoplasms, vascular
 CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence
 CC represents a human cell surface protease.
 XX
 SQ Sequence 688 AA:
 Alignment Scores:
 Pred. No.: 1,67e-178 Length: 688
 Score: 2963.00 Matches: 554
 Percent Similarity: 99.6% Conservative: 2
 Best Local Similarity: 99.3% Mismatches: 2
 Query Match: 88.9% Indels: 0
 DB: Gaps: 0
 US-10-806-370-11 (1-1748) x AD110395 (1-688)
 QY 11 ATGAGAGGGGACAGCGGGAATGCATCTCCAGCAAGAACCTTCAGCTGAGCATCT 70
 DB 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAAGCCAGGCATCTCCAGCTGGGACACTCCAGCCCGGGCATCTCCAGCCAGGCATCT 130
 DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAAGCCAGGCATCTCCAGCTGGGACACTCCAGCCCGGGCATCTCCAGCCAGGCATCT 190
 DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCTGTACACTCCAGCGCGGGCATCTCCAGCGGGGATCTCCAGCCAGGCATCT 250
 DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAAGCCAGGCATCTCCAGCTGGGACACTTCAGTTCAGTCTCCAGCGGCATCTCA 310
 DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
 QY 311 TCCGCGAGGTCAAGCTGGGTGAGCAACCTCCCAACAGATGTACTGTTAGAGCAACA 370
 DB 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValIleValAlaGlyAlaThr 120
 QY 371 CCAAGTGGGGGCTGTACCATTCGATCCGATCTCTGCGAGTACAGCACAGCAAGCGGCGC 430


```

Db      |||ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
Qy      431 ACCAGGGAGAGCCCGATACGAGCGCTGCGCAAGTTCACTCGCGGAGAGGCGCAAGAGC 490
Db      141 ThrArgIuSerProGlyThrSerIeuProIuSerIeuThrArgGlyGlyGlyIeu 160
Qy      491 CTACCGCATCGGGTGGCTGCTCTCTCAATTGCGCTGTGTGTTTCGCTCATCATCTCTC 550
Db      161 LeuProIeuIleGlyCysValIeuIeuIeuIleAlaIeuValIeuSerIeuIleIleIeu 180
Qy      551 TTCCAGTTCTGGCAGGCGCCACAGAGGATCAGGTACAGAGCAGAGAGGAGAGCTGTCC 610
Db      181 PheGlnPheTrpGlnGlyIleThrGlyIleArgTyrIleGlnArgGlyIuSerCysPro 200
Qy      611 AAGACCGCTGTTCGCTGAGCGGGGTGGTGGATGTCGACGAGCTGAGAGAGAGAGTGGCG 670
Db      201 LysHlValAlaValArgCysAspPoiIyAlaIAspCysIuSerIuSerAspPoiIuGly 220
Qy      671 TGCCTGAGGTTTGAAGTGGAGCAAGTCTGCTTAAATCTACTGAGGTCCTCCATCAG 730
Db      221 CysValArgPheAspTrpAspIuSerIeuIeuIuSerIeuIuSerGlySerSerHlSerIn 240
Qy      731 TGCCTTCCCATCTGTGAGCAGCACTGGAATGACTCTCTACTAGAGAGAGACCTGCCAGAG 790
Db      241 TrpIeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlyIuSerThrCysGlnGln 260
Qy      791 CTGGGTTTGAAGAGTCAACCGGCAACGAGGTTGGCCACAGGAGATTTGGCAACAGC 850
Db      261 LeuGlyPheGlyIuSerAlaHlArgThrArgIuValAlaHlArgAspPheAlaAsnSer 280
Qy      851 TTCTCAATCTTGAGATACAACTCCACATCCAGAGAAAGCTCCACAGGTTGAAATGCCCT 910
Db      281 PheSerIleIeuArgTyrAsnSerThrIleGlnIuSerIeuHlArgSerGlyIuSerPro 300
Qy      911 TCCACGCGGTATATATCTCCCTCAAGTGTCCCACTGCGGATGAGGCGCATGACCGGCGG 970
Db      301 SerGlnArgTyrIleSerIeuGlnCysSerHlCysGlyIuSerAlaIleMetThrGlyArg 320
Qy      971 ATCGTGGAGGGGGGCTGGCGCTCGGATAGCAAGTGGCTTGGCAAGTAGTGCACCTTC 1030
Db      321 IleValGlyGlyAlaIeuAlaSerAspSerIuSerTrpProTrpGlnValSerIeuHlPhe 340
Qy      1031 GGACACCAACCATCTGTGAGAGCAGCTCATTTGACGCCAGAGGAGTGTCTCACTGCCGC 1090
Db      341 GlyThrThrHlArgIleCysGlyGlyThrIeuIleAspAlaGlnTrpValIeuThrAlaIa 360
Qy      1091 CACTGCTTCTTCTGAGACCGGAGAGAGTCTGAGAGGCTGGAAGTGTACGGCGGACCC 1150
Db      361 HlCysPhePheValThrArgIuIuValIeuGlnGlyTrpIuValTyrAlaGlyThr 380
Qy      1151 AGCAACCTGACCAAGTGGCTGAGGAGAGCTCCATGGCGAGATCATATCAACAGCAT 1210
Db      381 SerIuIeuHlArgIuIuSerIeuProGlnAlaIleSerIleAlaGlnIleIleIleAsnSerIu 400
Qy      1211 TACACCGATGAGAGAGACTATGACATCGCCCTCATCGCGCTGTCCAAAGCCCTGACC 1270
Db      401 TyrThrArgIuGlnIuAspAspTyrAspIleAlaIeuMetArgIeuSerIuProIeuThr 420
Qy      1271 CTGTCCGCTCAATTCACCTGCTGTGCTCCCATGATGACAGACCTTTAGCCTCAAT 1330
Db      421 LeuSerAlaHlArgIleHlArgAlaCysIeuProMetHlArgGlnThrPheSerIeuAsn 440
Qy      1331 GAGACCTGCTGATACAGAGCTTGGCAAGACAGGAGAGAGATGAGAAACATCCCCC 1390
Db      441 GlnThrCysTrpIleThrGlyPheGlyIuSerThrArgIuIuTrpAspIuSerIeuThrSerPro 460
Qy      1391 TTCTCCCGAGAGTGCAGATCAATCTCACTCAAGATTCAGAAATGCAATGACTTGTGTC 1450
Db      461 PheIeuArgIuValGlnValAlaIeuIleAspPheIuSerCysAsnAspTyrIleuVal 480
Qy      1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGACCTTGTGGGCGCAGAGAC 1510

```

```

Db      481 TyrAspSerTyrIleuThrProArgMetMetCysAlaGlyAspLeuArgGlyIuArgAsp 500
Qy      1511 TCTTGCCAGGGAGACAGCGGGGGGCGCTTGTCTGTGAGCAGAAACCGCTGGTACCTG 1570
Db      501 SerCysGlnGlyAspSerGlyIuProIeuValCysGlnGlnIuAsnAsnArgTrpIuIeu 520
Qy      1571 GCAGGTGTACACAGCTGGGAGCAGGCTGTGGCCAGAGAAACAAACCTGTGTGTACACC 1630
Db      521 AlaGlyValThrSerTrpGlyThrGlyCysGlyIuArgIuAsnIuSerProGlyValTyrThr 540
Qy      1631 AAAGTACAGAAAGTTCTTCTCGATTTTACAGCAAGATGAGAGACGAGTCCGA 1684
Db      541 LysValThrGlyIuValIeuProTrpIleTyrSerIuSerMetGluAsnArgAlaGln 558

RESULT 9
ADJ46919
ID      ADJ46919 standard; protein; 688 AA.
AC      ADJ46919;
XX      06-MAY-2004 (first entry)
DT      XX
DE      Human transmembrane serine protease (MTSP)-related polypeptide #3.
KW      Human; transmembrane serine protease; MTSP; cell surface protease;
KW      plasmin; prostate specific antigen; PSA; proliferative disease;
KW      cell-surface protease-associated disease; autoimmune disease;
KW      inflammatory disease; infectious disease; endocrine disease; cancer;
KW      ocular disorder; cardiovascular disorder; chronic inflammatory disease;
KW      wound; circulatory disorder; dermatological disorder; restenosis;
KW      rheumatoid arthritis; psoriasis; diabetes; diabetic retinopathy;
KW      laser surgery scarring; glaucoma filtering surgery scarring;
KW      macular degeneration; CRBSF syndrome; bacterial infection; viral disease;
KW      solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
KW      Kaposi's sarcoma; enzyme.
XX      Homo sapiens.
OS      US2004001801-A1.
PN      01-JAN-2004.
PD      23-MAY-2002; 2002US-00156214.
PE      23-MAY-2002; 2002US-00156214.
PR      23-MAY-2002; 2002US-00156214.
PA      (CORV-) CORVAS INT INC.
PI      Madison BL, Sempke JF, Vlasuk GP, Kemp SJ, Komandla M, Slev DV;
XX      MPI; 2004-190126/18.
DR      N-PSDB; ADJ46918.
PT      Conjugate useful for treating e.g. cancer, cell-surface protease-
PT      associated diseases, comprising a peptidic substrate or nucleic acid
PT      substrate linked to a therapeutic agent through a linker.
XX      Claim 9; SEQ ID NO 26; 361pp; English.
XX
XX      The invention relates to a conjugate comprising a therapeutic agent and a
XX      peptidic substrate or nucleic acid substrate linked to the agent
XX      optionally through a linker or peptidic linker, where the peptidic
XX      substrate is proteolytically cleaved by a cell surface protease or a
XX      soluble, released or shed form conjugate to liberate the agent and the
XX      conjugate is not substantially cleaved by plasmin or prostate specific
XX      antigen (PSA). The conjugate is useful for treating a disease, which
XX      involves administering a conjugate to a subject, where the disease is
XX      preferably a proliferative disease or a cell-surface protease-associated
XX      disease. The diseases include autoimmune diseases, inflammatory diseases,
XX      infectious diseases and endocrine diseases. The conjugate is useful for
XX      treating a cell-surface protease-associated disease, which involves
XX      administering a conjugate comprising an agent and a peptidic substrate to
XX      a subject exhibiting symptoms of a cell-surface protease-associated

```


CC disorder, where the disease is selected from cancer, ocular disorders,
 CC cardiovascular disorders, chronic inflammatory diseases, wounds,
 CC circulatory disorders, dermatological disorders, resenosis, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, scarring from laser
 CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
 CC CHEST syndrome, bacterial infections, viral diseases, solid neoplasms and
 CC vascular tumors such as lung cancer, colon cancer, prostate cancer,
 CC melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
 CC serine protease (MSP)-related polypeptide of the invention.

XX Sequence 688 AA:

Alignment Scores:

Pred. No.:	1.67e-178	Length:	688
Score:	2963.00	Matches:	554
Percent Similarity:	99.6%	Conservative:	2
Best Local Similarity:	99.3%	Mismatches:	2
Query Match:	88.9%	Indels:	0
		Gaps:	0

US-10-806-370-11 (1-1748) x ADV46919 (1-688)

```

QY 11 ATGAGAGGAGACACCGGAGATGATCTCCAGCAAGAACCTTCACTGAGCATCT 70
DB 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCCAAGCATCTCCAGCTGGAGACACTCCAGCGCGGATCTCCAGCCAGCATCT 130
DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCCAAGCATCTCCAGCTGGAGACACTCCAGCGCGGATCTCCAGCCAGCATCT 190
DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGGTACACCTCCAGCGCGGATCTCCAGCGCGGATCTCCAGCCAGCATCT 250
DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAGCGCGGATCTCCAGCTGGAGACACTCCAGCTGGATCTCCAGCGGATCTCA 310
DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
QY 311 TCCGCCAGCTGAGCTCGGTGACAAACCTCCCAACAGAGTACTTGTAGAGCAACA 370
DB 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCACTGGGGGCTGTACCATCCATCATCTCTGCGAGTCCAGACCAACAGAGGCC 430
DB 121 ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProIleThrArgAla 140
QY 431 ACCAGGAGAGACCCAGTACAGACCTGCGCAAGTTCACTTGGCGGAGGCCAAGACAG 490
DB 141 ThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlnGlnSerGln 160
QY 491 CTACCGCTCAACGGGTGGGTGCTCTCTCTCAATGGCCCTGGGTGTTCCCTATATCTCT 550
DB 161 LeuProLeuIleGlyCysValLeuLeuIleAlaLeuValValSerLeuIleLeuLeu 180
QY 551 TTCAGTCTTGGAGGCGCACAGAGGATCAGTACAGAGCAGAGGAGAGGAGTGTCTCC 610
DB 181 PheGlnPheTrpGlnGlyHisThrGlyIleArgTyrIleGlnGlnArgGlnSerCysPro 200
QY 611 AAGCAGCGTGTTCGTGTGACCGGGGTGTGACTGCAAGCTGAAGAGTGAAGCTGGGC 670
DB 201 LysHisAlaValAlaCysAspGlyValAlaAspCysLysLeuLysSerAspGlnLeuGly 220
QY 671 TGGGTGAGGTTTGAAGTGGAGCAAGTCTGTTAAATCTACTCGGGTCCCTCCATCAG 730
DB 221 CysValAlaArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHisGln 240
QY 731 TGGCTTCCCATCTGAGAGCAACTGAGATGACTCTACTAGAGAGAACCTGCGCAGAG 790
DB 241 TrpLeuProIleCysSerSerSerAsnTrpAsnAspSerTyrSerGlnLysThrCysGlnGln 260

```

```

QY 791 CTGGGTTTGAGAGTGTCTACCGGACCAACGAGGTTGCCACAGGATTTGCCAAGC 850
DB 261 LeuGlyPheGlnSerAlaHisAlaGlyThrThGlyValAlaHisArgAspPheAlaSerSer 280
QY 851 TTCATCACTTGGAGATCAACCTCCACCATCCAGAAAGCTCCACAGGTGTGAATGCCCT 910
DB 281 PheSerIleLeuArgTyrIleAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCysPro 300
QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCCTCCAGTCCGAGGATGAGGGCATGACCGGCGG 970
DB 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGAGGCGGCTCGGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTCTCACTTC 1030
DB 321 IleValIleValAlaLeuAlaSerSerLysTrpProTrpGlnValSerLeuHisPhe 340
QY 1031 GGCACCAACCAATCTGTGAGAGCGCTCATTTAGCCCCAGTGGTGTCTACTGCGCGC 1090
DB 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTTCTTTCGTGACCCGGGAGAGAGTCTTGAGGGCTGGAAAGGTATCGCGGCACC 1150
DB 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACCTGACACCGTTCGAGGAGCGCTCCCATTTGCCGAGATCATCAACAGCAAT 1210
DB 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleAsnSerAsn 400
QY 1211 TACACGATGAGAGAGAGCATATGACATGCGCTCCATGCGGCTGTCCAAAGCCCTGACC 1270
DB 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY 1271 CTGTGCTGATCATCCACCTGCTTGCCTCCCATGATGACAGACCTTACCTCAAT 1330
DB 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGCGAGATCAACAGCTTTGGCAAGACCGGAGACAGATGACAGACATCCCCC 1390
DB 441 GluThrCysTrpIleTrpGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
QY 1391 TTCCTCCGAGAGGTGACAGTCAATCTCATGACACTTCAAGAAATGCAATGACTTGGTTC 1450
DB 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACAGTTACTTACCCCAAGATATGTGTGCTGGGACCTTCTGCGGGCGACAGAC 1510
DB 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgGlyArgAsp 500
QY 1511 TCCCTGCAAGGAGACAGCGGGGGGCTCTGTGTGTGAGCAAGACCGCTGTACCTG 1570
DB 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY 1571 GCAGGTGTCAACAGCTGGGCGACAGGCTGTGGCCAGAGAAACAACTGTGTGTACACC 1630
DB 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY 1631 AAAGTGAAGAAAGTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGGTCCGA 1684
DB 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlnAsnArgAlaGln 558

```

RESULT 10

ADVS0054 ID ADVS0054 standard, protein, 688 AA.

ADVS0054;

19-MAY-2005 (first entry)

Human endothelase-2L.

Antiarthritic; Antirheumatic; Antiosteoporotic; Antidiabetic;
 Ophthalmological; Cardiovascular-Gen.; Vulnereary; Antiinflammatory;

KM Vascotropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
 KM rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
 KM dermatological disease; cancer; neoplasm; endothelialase-2; enzyme.
 OS Homo sapiens.
 XX WO2005019270-A2.
 PN 03-MAR-2005.
 PD 12-AUG-2004; 2004WO-US026148.
 PF 14-AUG-2003; 2003US-0495005P.
 PR 14-NOV-2003; 2003US-0520164P.
 XX (DYAX-) DYAX CORP.
 PA (DENB-) DENBREON CORP.
 XX Madison EL, Nixon A;
 PI WPI; 2005-202609/21.
 DR N-PSDB; ADY50053.
 PT Novel protein capable of inhibiting endothelialase-2, useful for treating
 PT or preventing angiogenesis related disorder e.g. cancer.
 XX Disclosure; SEQ ID NO 2; 157bp; English.
 XX The invention relates to an isolated protein (I) which comprises a heavy
 CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
 CC immunoglobulin variable domain sequence, where the first and second
 CC immunoglobulin variable domain sequences from an antigen binding site
 CC that specifically binds to human endothelialase-2. (I) is useful for
 CC detecting an endothelialase or endothelialase activity in a sample, for
 CC modulating an activity of an ET2-expressing cell, for modulating
 CC proteolysis, for killing or inhibiting growth of a cell, for detecting
 CC endothelialase in a subject, for modulating endothelialase activity in a
 CC subject, for treating or preventing a disorder characterized by unwanted
 CC angiogenesis in a subject. The disorder is chosen from rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
 CC pterygii recurrence, scarring excimer laser surgery and glaucoma
 CC filtering surgery, cardiovascular disorders, chronic inflammatory
 CC disorders, wound repair, circulatory disorders, crest syndromes,
 CC dermatological disorders and cancers. The present sequence represents the
 CC human endothelialase-2L.
 XX
 SQ Sequence 688 AA;
 Alignment Scores:
 Pred. No.: 1.67e-178 Length: 688
 Score: 2963.00 Matches: 554
 Percent Similarity: 99.6% Conservative: 2
 Best Local Similarity: 99.3% Mismatches: 2
 Query Match: 88.9% Indels: 0
 DB: 9 Gaps: 0
 US-10-806-370-11 (1-1748) x ADY50054 (1-688)
 QY 11 ATGGAGAGGAGCAGCAGCGGAGATGATCTCCAGCAAGAACCTTGAGTGAGCATCT 70
 Db 1 MetcunargabserhlsaglyshmlaserProhlaahgthrProserhlaaglyhlsaser 20
 QY 71 CCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGCGCGGAGCATCTCCAGCCAGGACATCT 130
 Db 21 ProhlaaglnalaserProhlaaglyThrProhloglyhargalaserProhlaaglnalaser 40
 QY 131 CCAGCCCAAGGATCTCCAGCTGGAGACCTTCAGCGCGGAGCATCTCCAGCCAGGACATCT 190
 Db 41 ProhlaaglnalaserProhlaaglyThrProhloglyhargalaserProhlaaglnalaser 60
 QY 191 CCAGCTGGTACACCTCCAGCGCGGAGCATCTCCAGCGCGGAGCATCTCCAGCCAGGACATCT 250
 Db 61 ProhlaaglyhthrProhloglyhargalaserProhlaaglnalaserProhlaaglnalaser 80

QY 251 CCAGCCCGGAGCATCTCCGCGCTTGAGCATCACTTCCAGCTCTCATCCGAGGATCA 310
 Db 81 ProhlaarghalsaserProhlaaleuhalaserleuserhargsserserGlyhrgsser 100
 QY 311 TCCGCCAGGTCAGCCTGGGTGACCACTCCGCCAACAGAGTACTCTTTAGAGCAACA 370
 Db 101 SerlaargsserhalsaserValThrThrserProhargValThrLeuValargAlathr 120
 QY 371 CCAGTGGGGGCTGATCCATCCGATCTCTCCAGCTGAGGACAGCAACAGAGGCC 430
 Db 121 ProvalgllyalavalProhleargsserProhlaargsserhlaProhlaThargla 140
 QY 431 ACCAGGAGAGCCCAAGGATACAGCCTGCCAAGTTCACTGGCGGAGGCGCAGAACGAG 490
 Db 141 ThragluserProhloglyThrserserProhysrhetrThraglgllyGlnlyGln 160
 QY 491 CTACCGCTCATCGGGTGGCTGCTCTCTCTCATTTGCCCTGGTGGTCTCATCTCTC 550
 Db 161 LeuProleuilegilyCyvalLeuLeuLeuLeuLeuValaleuValaserLeuileleu 180
 QY 551 TTCAGTTCTGGCAGGGCCACACAGGATCAGTACAAGGACAGAGGAGAGAGCTGCC 610
 Db 181 PheglhphetrpglnghlyhsthnglyllehrglyrlysglnghargluserCysePro 200
 QY 611 AAGCAGCCTGTTCCGCTGTGACCGGGGTGTGTGACTGCAAGCTGAAGATGACAGCTGGGC 670
 Db 201 LysHlsalavalArgCyasrpglyValalAsrCyelsleuLysaserGlnleuGly 220
 QY 671 TGCGTGAAGTTTGACTGGGACAACTCTCTGTTAAATCTACTCTGGGTCTCTCCATCAG 730
 Db 221 CyvalargPrehaptPAspLysSerleuLysleLysrGlyseSerhlgln 240
 QY 731 TGCGTCTCCATCTGTACAGCACTGGAATGACTCTACTACAGAGAACCTGSCACAG 790
 Db 241 TrleuProlleCysserserhntPAspserlyserGlnlyThrCyglngln 260
 QY 791 CTGGGTTTGAAGTGTCAACCGGACCAACCGAGTTGCCACAGGATTTTGCACAGAC 850
 Db 261 LeuGlyhegluserhlaHlsargThrThnglyValalHlsargAsrPhehlaAsnser 280
 QY 851 TTCTCATCTTTGAGATACAACTCCACCATCCAGGAAGGCTCCCAAGTGTGAATGCCCT 910
 Db 281 PheSerlleuarglyrAsnserThrlleGlnluserleuHlsargSerGlnCysePro 300
 QY 911 TCCAGCGGATATATCTCCCTCCAGTGTTCACATCGCGGACTGAGGACATGACCGGCGG 970
 Db 301 SerGlnhrglyrllserserleuGlnCyseHlsCyseGlyleuargAlametThrglyhrg 320
 QY 971 ATCGTGGAGGGGCGCTGGCCTCGGATAGCAAGTGGCTTGCAAGTGAAGTTCACATTG 1030
 Db 321 llevalgllyllylaleuhalaserAsrserlyetrProhrglnvalaserleuHlsPhe 340
 QY 1031 GGCACCAACCAACTCTGTGAGGACAGCTCATTTGACCCAGTGGGGTGCATCTGCCGCC 1090
 Db 341 GlyThrThrhlsleCyseGlyGlyThrleuilePhehlaaglntrPvalleuThhlaala 360
 QY 1091 CACTGCTTTCTTGTGACCGCGGAGAAAGTCTGTGAAGGCTGGAAGGATGACGCGGACCC 1150
 Db 361 HlsCysePhehelaThrhargGlnlyvalleuGlnlytrPlyseValThrAlaaglyThr 380
 QY 1211 TACACCGATAGAGAGAGCACTATGACATGGCCCTCATGGGGCTGTCCAGCCCTTACC 1270
 Db 401 TyrThrhspgluglnAspAspTyraAspillealeuLeuclargleuSerlyseProleuThr 420
 QY 1271 CTGTCCGCTCACATCCACCTGCTGGCTCCCATGATGACAGACCTTTAGGCTCAAT 1330
 Db 421 LeuSerhlaHlsleHlsProlhlaCyseleuPromethlsGlyGlnlntrPheSerleuAsn 440

```

QY 1331 GAGACCTGCTGATCAGAGGCTTTGGCAAGACAGAGGAGACAGATGACAAAGATCCCCC 1390
DB 441 GlnThrCysTrpIleThnGlyPheGlyLysThrArgIleThrAspLysThrSerPro 460
QY 1391 TTCTCCGGGAGGTGACAGGTATCTCATCTGACATTCAGAAATGCAATGACTTGGTTC 1450
DB 461 PheLeuArgIleValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACGTTACCTTACCCCAAGATGATGCTGGGACCTTCCTGGGGGAGAGAC 1510
DB 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuArgIleValArgAsp 500
QY 1511 TCCTGACGAGGAGACAGCGGGGCGCTCTTGTCTGTGAGCAGAACACCGCTGGTACCTG 1570
DB 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTyrIleu 520
QY 1571 GCAGGTGTCACAGCAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGTGTGTACAC 1630
DB 521 AlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY 1631 AAAGTGCAGAGAAGTCTTCCCTGATTTACGCAAGATGAGAGCGAGGTCCGA 1684
DB 541 LysValThrGlnValLeuProTyrIleTyrSerLysMetGluAsnArgAlaGln 558

RESULT 11
ADE31743 ID ADE31743 standard; protein: 581 AA.
AC ADE31743;
XX
XX 29-JAN-2004 (first entry)
DT
XX
DE Human 2914 protein #SEQ ID 100.
XX
XX Antitreticlosclerotic; cardiant; vasotrophic; antiinflammatory;
XX thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
XX cardiovascular; disorder; ischemia; aortic bending;
XX KM vascular heart disease; endocarditis; atrial fibrillation; heart failure;
XX angina; cardiomyopathy; cardiac death.
XX
XX Homo sapiens.
XX
XX MO2003065984-A2.
XX
XX 14-AUG-2003.
XX
XX 29-JAN-2003; 2003WO-US002571.
XX
XX 01-FEB-2002; 2002US-0353224P.
XX 15-MAR-2002; 2002US-0364529P.
XX 19-APR-2002; 2002US-0373861P.
XX 29-APR-2002; 2002US-0376287P.
XX 12-JUN-2002; 2002US-0388080P.
XX 24-JUN-2002; 2002US-0390971P.
XX 03-JUL-2002; 2002US-0394130P.
XX 10-JUL-2002; 2002US-0394797P.
XX 21-AUG-2002; 2002US-0404904P.
XX 23-AUG-2002; 2002US-0405450P.
XX 04-SEP-2002; 2002US-0408707P.
XX 06-NOV-2002; 2002US-0424300P.
XX 05-DEC-2002; 2002US-0431042P.
XX 05-DEC-2002; 2002US-0431079P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Logan TV, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
XX Scagliano N, Perodin J, Rodrigue-Way A;
XX WPI; 2003-731468/69.
XX
XX N-PSDB; ADE31742.
XX
XX Identifying a compound capable of treating a cardiovascular disorder
XX (e.g. atherosclerosis) comprises assaying the ability of the compound to
XX

```

```

PT modulate the expression or activity of e.g. 1682, 6169 or 6193
PT polypeptide or nucleic acid.
XX
XX PS Disclosure; SEQ ID NO 100; 328pp; English.
XX
XX The invention relates to a method for identifying a compound capable of
XX treating a cardiovascular disorder. The present invention identifies the
XX differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
XX 43726, 69292, 21656, 32427, 2402, 7747, 1270, 9151, 60491, 1371, 7077,
XX 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 33394, 3484,
XX 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
XX 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
XX 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3454, 9792, 15400, 1452 or
XX 6585 genes in cardiovascular disease states. The methods are useful in
XX diagnosing, preventing and treating cardiovascular disorders, such as
XX atherosclerosis, cardiac hypertrophy, ischemia, reperfusion injury,
XX stenosis, arterial inflammation, vascular wall remodeling, coronary
XX microembolism, tachycardia, bradycardia, pressure overload, aortic
XX bending, coronary artery ligation, vascular heart disease, valvular
XX disease, including but not limited to, valvular degeneration caused by
XX calcification, rheumatic heart disease, endocarditis, or complications of
XX artificial valves; atrial fibrillation, long-QT syndrome, congestive
XX heart failure, sinus node dysfunction, angina, heart failure,
XX hypertension, atrial fibrillation, atrial flutter, pericardial disease,
XX including but not limited to, pericardial effusion and pericarditis;
XX cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
XX cardiomyopathy, myocardial infarction, coronary artery disease, coronary
XX artery spasm, ischemic disease, arrhythmia, sudden cardiac death and
XX cardiovascular developmental disorders. The methods may also be used for
XX identifying compounds that modulate cardiovascular disorders. Sequences
XX given in ADE31644-ADE31769 represent the genes and proteins that may be
XX regulated by a compound of the invention.
XX
XX SQ Sequence 581 AA:
XX
XX Alignment Scores:
XX Pred. No.: 4,49e-178 Length: 581
XX Score: 2956.00 Matches: 553
XX Percent Similarity: 100.0% Conservative: 1
XX Best Local Similarity: 99.8% Mismatches: 0
XX Query Match: 88.7% Indels: 7
XX DB: Gaps: 0

US-10-806-370-11 (1-1748) x ADE31743 (1-581)
QY 11 ATGAGAGGAGCAGAGCGGAATGATCTTCAGCAAGAACACCTTACCTGAGCATCT 70
DB 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGGGCGATCTCCAGCCGAGCATCT 130
DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCGAGCATCTCCAGCTGGGACACCTCCGAGCGGGGCGATCTCCAGCCGAGCATCT 190
DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGTACACCTCCAGGCGGGGCGATCTCCAGGCGGGGCGATCTCCAGCCGAGCATCT 250
DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAGCCGAGCATCTCCGAGCTGGGACACCTCCGAGCGGGGCGATCTCCAGCCGAGCATCT 310
DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerSerSerSerGlyArgSerSer 100
QY 311 TCAGCCGAGCTCAGCTGTGAGCAACCTCCCAACAGAGTGTACTTGTAGAGCAACA 370
DB 101 SerAlaArgSerAlaSerValThrThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAGTGGGGGCTGTACCATCCGATCATCTCTGCGCCAGGTGAGCAGCAGCAACGAGGCC 430
DB 121 ProValGlyValAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140

```

```

QY 431 ACCAGGAGAGAGCCAGGTACAGAGCTGCCAAGTTCACTGGCGGAGGCGCAGAAAGCAG 490
    |||||
Db 141 ThrArgIuSerProGlyThrSerLeuProLysPheThrTrpArgGluGlyGlnGln 160
QY 491 CTACCGGTCACTCGGGGTGGCTCTCTCTCATGGCCCTGGGAGTTCCTCATCATCTCTC 550
    |||||
Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValIleSerLeuIleIleLeu 180
QY 551 TTCCAGTTCTGGCAGGCGCAGACAGAGATCAGGTACAGAGAGAGAGAGAGAGCTGTCCC 610
    |||||
Db 181 PheGlnPheTrpGlnGlyIleThrGlyIleArgHlyLysGlnGlnArgGlnSerCysPro 200
QY 611 AAGCAGCTGTCTCCGTGTAGCGGGGTGTGAGCTGCACCAAGCTGAGAGAGTGAAGCTGGCC 670
    |||||
Db 201 LysHISaIaValArgCysAspGlyValIaAspCysLysLeuLysSerAspGlnLeuGly 220
QY 671 TGGGTGAAGTTGAGTGGGACAAAGTCTGTAAATCTACTCTGGGTCTCCCATAG 730
    |||||
Db 221 CysValArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHISgln 240
QY 731 TGGCTTCCCATCTGTAGCAGAGCACTGGAATGACTCTACTAGAGAGAGAGCTGCAGAG 790
    |||||
Db 241 TrpLeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlnLysThrCysGlnGln 260
QY 791 CTGGGTTTCCAGAGATGCTCACCGGACCAACGAGTTGCCACAGAGATTTTCCACAGAC 850
    |||||
Db 261 LeuGlyPheGlnSerAlaHISArgThrGlyValAlaHISArgAspPheAlaAsnSer 280
QY 851 TTTCATCTTGAGATCAACTCCACCATCCAGAGAAAGCTCCACAGGTCTGAATGAGCTT 910
    |||||
Db 281 PheSerIleLeuAlaGlyTyrAsnSerThrIleGlnGlnSerLeuHISArgSerGlnCysPro 300
QY 911 TCCGAGCGGATATCTCCCTCCAGTGTCCCACTGCGGAGCTGAGGCGCATGACCGGCGG 970
    |||||
Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHISGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGAGGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGAAGTTCGACTTC 1030
    |||||
Db 321 IleValGlyGlyAlaLeuAlaSerAspSerTyrTrpGlnValSerLeuHISpHe 340
QY 1031 GGCACACCCACATCTGTAGAGGAGCGCTCATAGACCCGAGGGGTCTCACTGCGCGCC 1090
    |||||
Db 341 GlyThrThrHISIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTCGTGACCCCGGAGAGAGCTCTGAGAGGCTGGAAGGTGTACGCGGAGACC 1150
    |||||
Db 361 HIEcysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACTGTGACCGAGTGGCTGAGGAGCGCTCCATTGGCCAGATCATATACAGCAAT 1210
    |||||
Db 381 SerAsnLeuHISglnLeuProGlnAlaAlaSerIleAlaGluIleIleAsnSerAsn 400
QY 1211 TACACCGATGAGAGAGCACTATGACATGAGCCCTCATGGCGGCTGTCCAGGCCCTGACC 1270
    |||||
Db 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuLeuArgLysSerLysProLeuThr 420
QY 1271 CTGTCCGTCACATCCACCTGCTTGGCTCCCATGATGAGACAGACTTTAGCCTCAAT 1330
    |||||
Db 421 LeuSerAlaHISIleHISProAlaCysLeuProMetHISglnGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGTGATGATCAAGGCTTTGGCAAGACGAGAGACATGACAGACATCCCCC 1390
    |||||
Db 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
QY 1391 TTCCCTCCGAGAGGTGAGGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTGATC 1450
    |||||
Db 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACAGTTACCTTACCCCAAGAGATGATGTGTCTGGGAGACTTTCGTGGGCGCAGAGAC 1510
    |||||
Db 481 TyrAspSerTyrLeuThrTrpAsnMetMetCysAlaGlyAspLeuValGlyGlyArgAsp 500
QY 1511 TCTCGCAGAGGAGAGACGCGGGGCGCTTGTGTCTGTAGACAGAAACCGCTGATACCTG 1570

```

```

Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY 1571 GCAGGTGACACCACTGGGAGCAGGCTGTGGCCAGAGAAACAAACCTGGTGTACAGC 1630
    |||||
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAlaLeuLysProGlyValTyrThr 540
QY 1631 AAAGTGAAGAGATTCTTCCCTGGATTTACAGCAAGATGAG 1672
    |||||
Db 541 LysValThrGluValLeuProTrpIleTyrSerLysMetGln 554

RESULT 12
ADH17430
ID ADH17430 standard; protein; 593 AA.
XX
AC ADH17430;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human NOV121 protein - SEQ ID 120.
XX
KW NOVX; antidiabetic; anorectic; cardiatic; hypotensive;
KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
KW prototoacide; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KW dermatological; antiaesthetic; antilipemic; metabolic; diabetes;
KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
KW epilepsy; immune; osteoarthritis; haemopoietic;
KW inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
KW cell differentiation; proliferation; haemopoiesis; wound healing;
KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
KW pharmacogenomic; human.
XX
OS Homo sapiens.
XX
PN W02003093432-A2.
XX
PD 13-NOV-2003.
XX
PF 02-MAY-2003; 2003WC-US013690.
XX
PR 02-MAY-2002; 2002US-0377321P.
PR 08-MAY-2002; 2002US-0378730P.
PR 24-MAY-2002; 2002US-0383075P.
PR 29-MAY-2002; 2002US-0384044P.
PR 30-MAY-2002; 2002US-0384215P.
PR 30-MAY-2002; 2002US-0384296P.
PR 30-MAY-2002; 2002US-0384297P.
PR 30-MAY-2002; 2002US-0384337P.
PR 30-MAY-2002; 2002US-0384352P.
PR 31-MAY-2002; 2002US-0385211P.
PR 02-JUL-2002; 2002US-0393333P.
PR 09-AUG-2002; 2002US-0402154P.
PR 09-AUG-2002; 2002US-0402171P.
PR 09-AUG-2002; 2002US-0402204P.
PR 09-AUG-2002; 2002US-0402205P.
PR 22-AUG-2002; 2002US-0405175P.
PR 27-AUG-2002; 2002US-0406129P.
PR 23-SEP-2002; 2002US-0412964P.
PR 30-SEP-2002; 2002US-0414975P.
PR 07-OCT-2002; 2002US-0416661P.
PR 24-OCT-2002; 2002US-0420851P.
PR 31-OCT-2002; 2002US-0422547P.
PR 01-MAY-2003; 2003US-00428275.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alvares E, Anderson DM, Boidog FL, Catterton E, Edinger SR,
PI Fernandes ER, Gerlach VL, Gorman L, Grose WM, Guo X, Ji W,
PI Kehuda R, Li L, MacDougall JR, Padigar M, Paturajan M,
PI Peterzon JD, Rastelli L, Shinkete RA, Spytek KA, Stone DJ,
PI Vermet CAM, Voss EZ, Zhong M;

```

XX WPI: 2004-053040/05.
DR N-PSDB: ADH17429.

PT New isolated NOXV polypeptide, useful for preventing, diagnosing or
PT treating NOXV-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

XX Claim 1, SEQ ID NO 120; 478bp; English.

XX The invention relates to a novel isolated NOXV polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiatic, hypotensive, antiarteriosclerotic, anorectic, vitricide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteoprotective, antidiabetic,
CC antiinflammatory, dermatological, antisthmatic and antilipemic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis. In gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOXV substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOXV protein of the invention.

XX Sequence 593 AA:

Alignment Scores:
Pred. No.: 6.49e-178 Length: 593
Score: 2953.50 Matches: 555
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 5
Query Match: 88.6% Indels: 0
DB: 8 Gaps: 1

US-10-806-370-11 (1-1748) x ADH17430 (1-593)

QY 8 ACATGAGAGAGGACAGCCAGGAGATCATCTCCAGCAGAGACCTTCAGTGGAGCA 67
DB 4 ThiretGluAgaBpSerHisGlyAsnHisSerProAlaArgThrProSerAlaGlyAla 23
QY 68 TCTCCAGCCAGGAGATCTCCAGCTGGGACACCTTCAGGCGGAGATCTCCAGCCAGGCA 127
DB 24 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 43
QY 128 TCTCCAGCCAGGAGATCTCCAGCTGGGACACCTTCAGGCGGAGATCTCCAGCCAGGCA 187
DB 44 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 63
QY 188 TCTCCAGCTGGTACACCTTCAGGCGGAGATCTCCAGGCGG-----GCA 232
DB 64 SerProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAla 83
QY 233 TCTCCAGCCAGGAGATCTCCAGGCGGAGATCTCCAGCTGGGATCTTCAGAGTCC 292
DB 84 SerProAlaGlnAlaSerProAlaArgAlaSerProAlaGlnAlaSerLeuSerArgSer 103
QY 293 TCATCCGCGAGAGTATCTCCAGGCGGAGATCTCCAGCTTCGAGCAACCTCCCAACAGAGATG 352
DB 104 SerSerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgVal 123
QY 353 TACCTTGTAGAGCAACAGAGTGGGAGCTGATCCCATCGATCATCTTCGACAGTCA 412
DB 124 TyrluValAlaArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSer 143

QY 413 GCACAGCAACAGAGGCGCACAGAGAGAGCCAGGTACAGAGCTTCCAGTTCACCTGG 472
DB 144 AlaProAlaThrThrArgAlaThrArgGluSerProGlyThrSerLeuProLysPheThrTrp 163
QY 473 CCGAGAGGCGCAGAGAGAGTACCGGCTCATCGGAGTGGCTCTCTTCATGGCCCTGGTG 532
DB 164 ArgGluGlyGlnLysGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuVal 183
QY 533 GTTTCGCTCATCATCTCTCCAGTTCCTGAGGCGCACAGAGATCCAGGTCAAGAG 592
DB 184 ValSerLeuIleLeuLeuPheGlnPheTrpGlnGlyHisThrGlyIleArgTyLysGlu 203
QY 593 CAGAGGAGAGAGTGTCCCAAGCAGCGCTGTGCTGACCGGAGTGGTGAATCAAGCTG 652
DB 204 GlnArgGluSerCysProLysHisAlaValArgCysAspGlyValAlaAspCysLysLeu 223
QY 653 AAGAGTACAGAGTGGAGTGGCTGAGAGTTCAGTGGAGCAAGTCTCTTAAATCTAC 712
DB 224 LysSerAspGluLeuGlyCysValArgPheAspTrpAspLysSerLeuLeuLysIleTy 243
QY 713 TCTGGGCTCTCCATCATGAGTGGCTTCCATGTGACAGCAATCGAATGACTCTACTCA 772
DB 244 SerGlySerSerHisGlnTrpLeuProIleCysSerSerAntTrpAspSerTySer 263
QY 773 GAGAGAGCTGCCAGAGAGTGGGTTTCAGAGTGTCAACGAGCAACCGAGTGGCCAC 832
DB 264 GlnLysThrCysGlnGlnLysLeuGlyPheGlnSerAlaHisArgThrThrGluValAlaHis 283
QY 833 AGGAGATTTTGCACCAAGCTTTCATCTTGAATACATCCACATCCAGAAAGCTTC 892
DB 284 ArgAspPheAlaAsnSerPheSerIleLeuValArgTrpAsnSerThrIleGlnGluSerLeu 303
QY 893 CACAGCTGTGAATGCCCTTCCACGCGGATATGTCCTCCAGTGTTCACCTGGAGCTG 952
DB 304 HisArgSerGluCysProSerGlnArgTyLysSerLeuGlnCysSerHisCysGlyLeu 323
QY 953 AGGSCCATGACCGGCGGAGATCTGGGAGGAGCGCTGGAGTGGAGTGGAGCTTGG 1012
DB 324 ArgAlaMetThrThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTrpProTrp 343
QY 1013 CAAGTAGTGTGACATTCGAGCAGCAGCAGCAGATGTGTGAGAGCAGCTATGACGCGCAG 1072
DB 344 GlnValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGln 363
QY 1073 TGGGTGCTCATCTGCGCGCCCATCTGTTTGTGTGACCCGGAGAAAGTCTGGAGGCTGG 1132
DB 364 TrpValLeuThrAlaAlaHisCysPhePheValThrArgGlnLysValLeuGluGlyTrp 383
QY 1133 AAGGTGTAACCGGCGCAGCAGCAGCAGCAGTGTGCTGAGGAGCGCTCATTTGCCGAG 1192
DB 384 LysValTyrlaGlyThrSerAsnLeuHisGlnLeuProGluAlaAlaSerIleAlaGln 403
QY 1193 ATCATCATCAACAGCATTTACACCGATGAGAGAGCAGTATGACATCGGCTTCATCGG 1252
DB 404 IleIleIleAsnSerAsnTyrlThrAspGluGluAspAspTyrlAspIleAlaLeuMetArg 423
QY 1253 CTGTCAAGCGCTGACCTGTGCTCGCTCAATCCACCTGCTTGGCTCCCATGATGAGA 1312
DB 424 LeuSerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGly 443
QY 1313 CAGACCTTTAGGCTCATGAGACCTGTGATGATGACAGAGCTTTGGCAGAGCAGGAGACA 1372
DB 444 GlnThrPheSerLeuAsnGlnThrCysTrpIleThrGlyPheGlyLysThrArgGlnThr 463
QY 1373 GATGACAAAGACATCCCTTCTCCGAGAGTGGAGGTCATCTCATGCAATTCAGAA 1432
DB 464 AspAspLysThrSerTrpPheLeuArgGluValGlnAlaMetLeuIleAspPheLys 483
QY 1433 TGCATGATCTATTGGTCTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGC 1492
DB 484 CysAsnAspTyrlLeuValTyrlAspSerTyrlLeuThrProArgMetMetCysAlaGlyAsp 503

QY 1493 CTTGTCGGGGCAGAGACTCTTCAGGAGAGACGGGGGCTTTGTCTGTAGACAG 1552
 DB 504 LeuArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGln 523
 QY 1553 AACACCGCTGTGCTCCTTCGCGAGGTGTACCCAGCTGGGGCAGAGGCTGTGGCCAGAGAAC 1612
 DB 524 AsnAsnArgTyrPheLeuAlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsn 543
 QY 1613 AACCTGCTGTGTACACCAAGTACAGAAAGTTCTTCCTCGATTTTACAGCAAGATGAG 1672
 DB 544 LysProGlyValTyrThrValThrGlyValLeuProTyrPheTyrSerLysMetGln 563

RESULT 13
 ADH17414
 ID ADH17414 standard; protein, 586 AA.
 XX ADH17414;
 AC
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human NOV12a protein - SEQ ID 104.

XX NOVX; antidiabetic; anorectic; cardiact; hypotensive;
 KM antidiabetic; anorectic; vitamin; antibacterial; fungicide;
 KM protocoarctide; nootropic; neuroprotective; antiparkinsonian;
 KM anticonvulsant; osteopathic; antiallergic; antiinflammatory;
 KM dermatological; antineoplastic; antihypertensive; metabolic; diabetes;
 KM obesity; infection; anorexia; cancer; cardiovascular; hypertension;
 KM atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
 KM epilepsy; immune; osteoarthritis; haemopoietic;
 KM inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
 KM cell differentiation; proliferation; haemopoiesis; wound healing;
 KM angiogenesis; gene therapy; chromosome mapping; tissue typing;
 KM pharmacogenomic; human.

XX
 OS Homo sapiens.
 XX
 PN WO2003093432-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 02-MAY-2003; 2003MO-US013690.

XX
 02-MAY-2002; 2002US-0377321P.
 PR 08-MAY-2002; 2002US-0378730P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384004P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384236P.
 PR 30-MAY-2002; 2002US-0384237P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 31-MAY-2002; 2002US-0385211P.
 PR 02-JUN-2002; 2002US-0383333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405175P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416611P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.

XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alvarez E, Anderson DW, Boldog FL, Catterton S, Edinger SR;
 PI Fernandes BR, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
 PI Kekuda R, Li L, Macdougall JR, Padigera M, Pattnajain M,
 PI Peterson JD, Rastelli L, Shimkels RA, Spletter KA, Stone DJ;

PI Vernet CAM, Voas EZ, Zhong M;
 XX
 DR WPI: 2004-053040/05.
 DR N-PSDB: ADH17413.
 DR
 XX
 PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
 PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
 PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 104; 478pp; English.

XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic,
 CC cardiact, hypotensive, antidiabetic, anorectic, vitamin,
 CC antibacterial, fungicide, protocoarctide, nootropic, neuroprotective,
 CC antiparkinsonian, anticonvulsant, osteopathic, antiallergic,
 CC antiinflammatory, dermatological, antineoplastic and antihypertensive
 CC activities. The polypeptide, nucleic acid molecules and antibodies may
 CC be useful in the manufacture of a medicament for treating metabolic
 CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
 CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
 CC diseases including hypertension and atherosclerosis, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
 CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
 CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
 CC acids and polypeptides may also be used as targets for the identification
 CC of small molecules that modulate or inhibit neurogenesis, cell
 CC differentiation, cell proliferation, haemopoiesis, wound healing and
 CC angiogenesis, in gene therapy and the in generation of antibodies that
 CC bind immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids may be further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine and pharmacogenomics. The current sequence is that of the human
 CC NOVX protein of the invention.

XX
 SO Sequence 586 AA;

Alignment Scores:
 Pred. No.: 1,34e-177 Length: 586
 Score: 2948.50 Matches: 554
 Percent Similarity: 99.1% Conservative: 0
 Best Local Similarity: 99.1% Mismatches: 0
 Query Match: 88.5% Indels: 1
 DB: 8 Gaps: 1

US-10-806-370-11 (1-1748) x ADH17414 (1-586)

QY 11 ATGAGAGGAGACAGCCAGCGGATGCTTCACAGAAAGACCTTACGTCGACATCT 70
 DB 1 MetClnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAAGCCAGGATCTCCAGCGGGGACACCTCCAGCGGCGATCTCCAGCCAGGATCT 130
 DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAAGCCAGGATCTCCAGCGGGGACACCTCCAGCGGCGATCTCCAGCCAGGATCT 190
 DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCGTACACCTCCAGCGGGGACATCTCCAGCGCGG-----GCATCT 235
 DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 236 CCAAGCCAGGATCTCCAGCGGGGACATCTCCAGCGGCGATCTCCAGGATCTCTCA 295
 DB 81 ProAlaGlnAlaSerProAlaGlyArgAlaSerProAlaGlnAlaSerLeuSerArgSer 100
 QY 296 TCCGCGAGTCAATCATCCGCGAGTCAAGCTCGGTGCAACCTCCACAGAGTGTAC 355
 DB 101 SerClyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
 QY 356 CTTGTTAGACAAACAGATGGGGGCTGTACCATTCGATCATCTTCTGCGAGTCA 415

Db 121 LeuValAlaArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
 QY 416 CCAGCACCAGGGGCGACAGGGAGAGCCCGAGTACAGCCCTGCGCCAGTTCACTGCGCG 475
 Db 141 ProIleAlaThrAlaArgAlaArgSerProGlyThrSerLeuProLysPheThrTyrArg 160
 QY 476 GAGGGCCAGAGACAGTACCGCTCATCGGGGCTGCTCTCTCATTTGGCCCTGGTGT 535
 Db 161 GluGlyGlnLysGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuVal 180
 QY 536 TCGCTCATCATCTCTTCCAGTTCTGGCAGGGCCACACAGGGATCAGGTACAGAGCAG 595
 Db 181 SerLeuIleIleLeuPheGlnPheThrGlnGlyIleThrGlyIleArgGlyLysGln 200
 QY 596 AGGAGAGCTGTCCCAAGCAGCTGTTCGCTGTGACGGGGGTGGAGCTGCAACTGAAG 655
 Db 201 ArgGlnSerCysProLysHsAlaValArgCysAspGlyValAlaAspCysLysLeuLys 220
 QY 656 AGTACAGAGCTGGGCTGGTGGAGTTGACTGGGACAAAGTCTTAAATCTACTACT 715
 Db 221 SerAspGlnLeuGlyCysValArgPheAspTyrAspLysSerLeuLeuLysIleTyrSer 240
 QY 716 GGGTCTCCATCATAGTGGCTTCCCATCTGTAGACAACTGGAATGACTCTACTCAAG 775
 Db 241 GlySerSerHisGlnTyrPleuProIleCysSerSerAsnTyrAsnAspSerTyrSerGln 260
 QY 776 AAGACCTGCCAGCGCTGGGTTTGGAGAGTCTCACCGGACACCGAGGTGGCCACAG 835
 Db 261 LysThrCysGlnGlnLeuGlyPheGlnSerAlaHisArgThrThrGlnValAlaHisArg 280
 QY 836 GATTTTGCACACAGCTTCTCAATCTTGAGATACAACTCCACACCCAGAAAGCTCCAC 895
 Db 281 AspheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHis 300
 QY 896 AGGTGTGAATGCCCTTCCACAGCGGTATATCTCCCTCAGTGTCCACTGCGGACTGAG 955
 Db 301 ArgSerGlnCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
 QY 956 GCCATGACCGGGCGGATCTGTGGAGGGCGCTGGCTGGATACAGTGGCTTGGCA 1015
 Db 321 AlaeThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTyrProTyrGln 340
 QY 1016 GTGAGTGTGCACTTCCGAGCACCACCACTCTGTGAGGCAAGCTCATTTGAGGCCAGG 1075
 Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuLysLeuAlaGlnTyr 360
 QY 1076 GTGCTCATCTGCGGCCACTGCTTCTTCTGTGACCGCGGAGAGTCTGAGGGCTGAG 1135
 Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgLysValLeuGlnGlyTyrLys 380
 QY 1136 GTTGTAGCGGGGACCAAGCACTGCAACCAAGTTGCTGAGGAGCGCTTCATTGCCGAGATC 1195
 Db 381 ValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIle 400
 QY 1196 ATCATCAACAGCAATTACACCGATGAGAGAGACACTTGTGACATCGCCCTCATCGGCTG 1255
 Db 401 IleLeuAsnSerAsnTyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeu 420
 QY 1256 TCCAGAGCCCTGAGCCCTGTCCGCTCACATCCACCTGCTTCCCTCCCATGATGAGACAG 1315
 Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
 QY 1316 ACCTTTAGCTCAATGAGACCTGTGTGATCACAGGCTTTGGCAAGACAGGAGACAGAT 1375
 Db 441 ThrPheSerLeuAsnGlnTyrThrCysTyrIleThrGlyPheGlyLysThrArgGlnThrAsp 460
 QY 1376 GACAAGACATCCCTCTCCCGGAGGTGAGGATCATCTCATGACTTCAAGAAATGC 1435
 Db 461 AspLysThrSerProPheLeuArgLysValAlaAsnLeuLysAspPheLysLysCys 480
 QY 1436 AATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGACCTT 1495
 Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeu 500

QY 1496 CGTGGGGGACAGACTCTCTCCAGGAGACAGCGGGGGCCCTTGTGTGTGACAGAAC 1555
 Db 501 ArgGlyGlyAlaArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsn 520
 QY 1556 AACCGCTGTACTGTGACAGGTGTCCACAGCTGGGGGACAGGCTGTGGCCAGAGAACAA 1615
 Db 521 AsnArgTyrTyrLeuAlaGlyValThrSerTyrGlyThrGlyCysGlnGlnArgAsnLys 540
 QY 1616 CCTGTGTGTACACCAAGTACAGAGAGTTCTTCCCTGGATTTACAGCAAGATGAG 1672
 Db 541 ProGlyValTyrThrLysValThrGlnValLeuProTyrPheTyrSerLysMetGln 559
 RESULT 14
 ADH17434
 ID ADH17434 standard; protein: 586 AA.
 AC ADH17434;
 DT 11-MAR-2004 (first entry)
 XX
 DE Human NOV12k protein - SEQ ID 124.
 XX
 KW NOX; antidiabetic; anorectic; cardiatic; hypotensive;
 KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
 KW protozoacide; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
 KW dermatological; antiaspatic; antilipemic; metabolic; diabetes;
 KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
 KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
 KW epilepsy; immune; osteoarthritis; haemopoietic;
 KW inflammatory skin disorder; asthma; dyplidemia; neurogenesis;
 KW cell differentiation; proliferation; haemopoiesis; wound healing;
 KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
 KW pharmacogenomic; human.
 OS Homo sapiens.
 XX
 PN M02003093432-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 02-MAY-2003; 2003WC-US013690.
 XX
 PR 02-MAY-2002; 2002US-0377321P.
 PR 08-MAY-2002; 2002US-0378730P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384044P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384286P.
 PR 30-MAY-2002; 2002US-0384287P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 30-MAY-2002; 2002US-0384352P.
 PR 31-MAY-2002; 2002US-0385211P.
 PR 02-JUL-2002; 2002US-0393333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405155P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416661P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alvarez E, Anderson DW, Boidog FL, Catterton E, Edinger SR,
 PI Fernandes ER, Gerlach VL, Gorman L, Grose WM, Guo X, Ji W,
 PI Kekuda R, Li L, Macdougall JR, Padigaru M, Patunrajan M,

PI Peterson JD, Rastelli L, Shimkets RA, Spytek KA, Stone DJ;
PI Vernet CAM, Voss EZ, Zhong M;
XX WPI: 2004-053040/05.
DR N-PSDB: ADH17433.
XX
PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
PS Claim 1; SEQ ID NO 124; 478bp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiatic, hypotensive, antiatherosclerotic, anorectic, virucide,
CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antirheumatic,
CC antiinflammatory, dermatological, antiasthmatic and antipileptic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis. In gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
XX
XX Sequence 586 AA;
SQ
Alignment Scores:
Pred. No.: 1.34e-177 Length: 586
Score: 2948.50 Matches: 554
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 5
Query Match: 88.5% Indels: 5
DB: 8 Gaps: 1
US-10-806-370-11 (1-1748) x ADH17434 (1-586)
QY 11 ATGAGAGGAGGACAGCCACGGGATGATCTTCACGAAGAACCTTGAGTGCATCT 70
Db 1 Mccglunrgraspserrh1eg1yAanhlaserProh1argrhnrProsserAlaG1yAlaser 20
QY 71 CCAAGCCAGGATCTTCAGCTGGGACACCTTCAGAGCCGGGATCTTCAGCCAGGATCT 130
Db 21 ProAlaGlnAlaSerProAlaG1yThrProProG1yArGAlaSerProAlaGlnAlaSer 40
QY 131 CCAAGCCAGGATCTTCAGCTGGGACACCTTCAGAGCCGGGATCTTCAGCCAGGATCT 190
Db 41 ProAlaGlnAlaSerProAlaG1yThrProProG1yArGAlaSerProAlaGlnAlaSer 60
QY 191 CCAAGCTGTACACTCCAGAGCCGGGACATCTTCAGAGCCGG-----GCATCT 235
Db 61 ProAlaG1yhrPrProG1yArGAlaSerProG1yArGAlaSerProAlaGlnAlaSer 80
QY 236 CCAAGCCAGGATCTTCAGAGCCGGGACATCTTCGGCTCTGGCATACATTCAGGATCTCA 295
Db 81 ProAlaGlnAlaSerProAlaG1yArGAlaSerProAlaGlnAlaSerLeuSerArGSerSer 100
QY 296 TCCGGAGGATCATATCCGACAGGTCAGGCTCGGTGACAACCTCCCAACGAGGTATAC 355
Db 101 SerG1yArGSerSerSerAlaArGSerAlaSerValThrThrSerProThnArGValTyr 120
QY 356 CTGTATTAGAGCAACCAAGTGGGGCTGTACCCATCCGATCATCTCTGCACAGTCA 415

Db 121 LeuValArGAlaThrnrProValG1yAlaValProIleArGSerSerProAlaArGSerAla 140
QY 416 CCAAGCAACAGGCGCCACCAAGGAGAGCCAGGTACAGACCTGCCCAAGTTACCTGGCG 475
Db 141 ProAlaThrArGAlaThrArGAlaSerProG1yThrSerLeuProLysPheThrTTPArG 160
QY 476 GAGAGCCAGAGACAGTACCGCTCATGGGTCGCTGCTCCATTCAGTCCCTGGTGGTT 535
Db 161 GlnG1yGlnSerGlnLeuProLeuIleG1yCysAlaLeuLeuLeuIleAlaLeuValAl 180
QY 536 TCGCTCATCATCTCTTCAGTTCTGGCAGAGCCACACAGGATCAGGTACAGAGACAG 595
Db 181 SerLeuIleIleLeuPheGlnPheTrpGlnG1yAlaThrG1yIleArG1yLeuG1yGln 200
QY 596 AGGAGAGGCTGTCCCAAGACAGCTGTTCCGTGTGACAGGGGTGTGACATCGAAGTAA 655
Db 201 ArGAlaSerCysProLysAlaAlaValArGCSArSpG1yAlaValArCybLysLeuLys 220
QY 656 AGTACAGAGCTGGCTGCGTAGAGTTGACTGGGACAACTCTGTAAATCTACTCT 715
Db 221 SerArSpG1yLeuG1yCysValArGPhaArPTTPArSpLysSerLeuLysIleTyrSer 240
QY 716 GGGTCTCCATCATAGTGGCTTCCATCTGTAGCAGCAACTGGAATGACTCTTACAGAG 775
Db 241 GlySerSerh1eg1nTrpLeuProIleCysSerSerArnrThraArSpArTyrSerGln 260
QY 776 AAGACCTGCCAGACAGCTGGTTTCAGAGAGTCTCACCGGCAACCGAGTTGCCCAAG 835
Db 261 LysThrCybGlnGlnLeuG1yPheGlnSerAlaAlaArGThrThrGlnValAlaAlaArG 280
QY 836 GATTTGCCAACAGCTCTCAATCTTGAGATPACAACTCCACCACTCAGAAAGCTCCAC 895
Db 281 AsPheAlaArSpArPheSerIleLeuArG1yThrArSpArThrIleGlnGlnSerLeuAla 300
QY 896 AGGTGTGAATGCCCTTCCAGCGGATATATCTCCCTCAAGTTTCCACTCGGACTGAG 955
Db 301 ArGSerG1yCysProSerGlnArG1yTyrIleSerLeuGlnCysSerh1eg1yLeuArG 320
QY 956 GCCATGACCGGGCGGATCGTGGGAGGGGCGCTGGCTCCGATAGCAAGTGGCTTGGCAA 1015
Db 321 AlArMetThrG1yArG1yAlaG1yG1yAlaLeuAlaSerArSpArTyrProTTPGln 340
QY 1016 GTGAGTGTGACACTGGGACACCCACATCTGTGAGAGCGCTCATTCAGCCCACTGG 1075
Db 341 ValSerLeuArh1ePheG1yThrThrAlaIleCysG1yG1yThrLeuIleArPAlaGlnTTP 360
QY 1076 GTGCTCATCTGCGGCCCATCTCTTGTGTGACCCGGGAGAAAGTCTGAGAGGCTGGA 1135
Db 361 ValLeuThrAlaAlaAlaAlaAlaCysPhePheValThrArG1yLysValLeuGlnG1yTTP 380
QY 1136 GTGTACCGGGGACACAGCAACCTGCACCACTGTCCTGAGCAGCTCCATTCGCGAATC 1195
Db 381 ValTyrAlaG1yThrSerArLeuAlaAlaGlnLeuProG1yAlaAlaSerIleAlaGlnIle 400
QY 1196 ATCATCAACAGCAATPACACCGATGAGAGAGAGACATATACATCGCCTCATGCGGCT 1255
Db 401 IleIleArSpArSerArnrTyrThrArPGLuGluArPArTyrArPAlaLeuMetArGlu 420
QY 1256 TCCAGGCCCTGACCTGTGCGCTCATCCATCCACTGCTTGCCTCCCATGATGACAG 1315
Db 421 SerTyrProLeuThrLeuSerAlaAlaIleAlaProAlaCysLeuProMetCh1eg1yGln 440
QY 1316 ACCTTTAGCTCATATGAGACCTGCTGATATACAGGCTTTGGCAAGACAGAGACAGAT 1375
Db 441 ThrPheSerLeuArSpArnrTyrTyrPAlaThrG1yPheG1yArG1yThrArG1yThrArP 460
QY 1376 GACAAGACATCCCGCTTCTCCGGAGAGTGCAGGTCATCTCATCGACTCAAGAAATGC 1435
Db 461 AsPArThrSerProPheLeuArG1yAlaGlnValAlaArLeuIleArPheLysCyb 480
QY 1436 AATGACTACTTGTGTATATGACAGTTACCTTACCAGATGATGTGTGCTGGGACCTT 1495

Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeu 500
QY 1496 CGTGGGGCGAGACTCTCTCCAGGAGAGAGGGGGGCTCTTGCCTTGACGAGAAC 1555
Db 501 ArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsn 520
QY 1556 AACCGGTGACTGTCGAGGTGTACACAGCTGGGGGACAGCTGTGGCCAGAGAACAA 1615
Db 521 AsnArgTyrTyrLeuAlaGlyValThrSerTyrGlyThrGlyGlyGlnArgAsnLys 540
QY 1616 CCGTGTGTGTACACCAAGTGTACAGAGAGTTCTTCCCTGATTTACAGCAAGATGAG 1672
Db 541 ProGlyValTyrThrLysValThrGlnValLeuProTyrLeuTyrSerLysMetGln 559
RESULT 15
ADH17452
ID ADH17452 standard; protein; 586 AA.
XX
XX ADH17452;
XX
XX 11-MAR-2004 (first entry)
XX
DE Human NOV12t protein - SEQ ID 142.
XX
XX NOVX; antidiabetic; anorectic; cardiatic; hypotensive;
XX antidiabetic; anorectic; anorectic; virucide; antibacterial; fungicide;
XX protocoacid; neuroprotective; antiparkinsonian;
XX anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
XX dermatological; antidiabetic; antidiabetic; metabolic; diabetes;
XX obesity; infection; anorexia; cancer; cardiovascular; hypertension;
XX atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
XX epilepsy; immune; osteoarthritis; haemopoietic;
XX inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
XX cell differentiation; proliferation; haemopoiesis; wound healing;
XX angiogenesis; gene therapy; chromosome mapping; tissue typing;
XX pharmacogenomic; human; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX MISC-difference 401
XX FT /note= "Wild-type Thr may be substituted by Ile due to
XX FT single nucleotide polymorphism (SNP)"
XX
XX
XX MO2003093432-A2.
XX
XX 13-NOV-2003.
XX
XX 02-MAY-2003; 2003MO-US013690.
XX
XX 02-MAY-2002; 2002US-0377321P.
XX 08-MAY-2002; 2002US-0378730P.
XX 24-MAY-2002; 2002US-0383075P.
XX 29-MAY-2002; 2002US-0384044P.
XX 30-MAY-2002; 2002US-0384215P.
XX 30-MAY-2002; 2002US-0384286P.
XX 30-MAY-2002; 2002US-0384287P.
XX 30-MAY-2002; 2002US-0384337P.
XX 30-MAY-2002; 2002US-0384352P.
XX 31-MAY-2002; 2002US-0385211P.
XX 02-JUL-2002; 2002US-0393333P.
XX 09-AUG-2002; 2002US-0402154P.
XX 09-AUG-2002; 2002US-0402171P.
XX 09-AUG-2002; 2002US-0402204P.
XX 09-AUG-2002; 2002US-0402205P.
XX 22-AUG-2002; 2002US-0405175P.
XX 27-AUG-2002; 2002US-0406129P.
XX 23-SEP-2002; 2002US-0412954P.
XX 30-SEP-2002; 2002US-0414975P.
XX 07-OCT-2002; 2002US-0416611P.
XX 24-OCT-2002; 2002US-0420851P.
XX 31-OCT-2002; 2002US-0422547P.
XX 01-MAY-2003; 2003US-00428275.

XX
PA (CURA-) CURAGEN CORP.
XX
XX Alvarez E, Anderson DW, Boldog FI, Catterson E, Edinger SR,
PI Fernandes ER, Gerlach VU, Gorman L, Grosse WM, Guo X, Ji W;
PI Kekuda R, Li L, MacDougall JR, Padigar M, Patirajan M,
PI Peterson JD, Raselli L, Shimkets RA, Splek KA, Stone DJ,
PI Vernat CM, Voss EZ, Zhong M;
XX
XX WPI; 2004-053040/05.
DR N-PSDB; ADH17451.
XX
XX
XX New isolated NOVX polypeptide, useful for preventing, diagnosing or
PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
XX
XX
XX Claim 1; SEQ ID NO 142; 478bp; English.
XX
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiatic, hypotensive, antidiabetic, anorectic, virucide,
CC antibacterial, fungicide, protocoacid, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
CC antiinflammatory, dermatological, antidiabetic and antidiabetic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
XX
XX
XX SQ Sequence 586 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,77e-177 Length: 586
XX Score: 2943.50 Matches: 553
XX Percent Similarity: 98.9% Conservative: 0
XX Best Local Similarity: 98.9% Mismatches: 1
XX Query Match: 88.3% Indels: 5
XX DB: 8 Gaps: 1
XX
XX
XX US-10-806-370-11 (1-1748) x ADH17452 (1-586)
QY 11 ATGGAGAGGAGCGAGCGAGGATGATCTCCAGCAAGACACTTCAGCTGAGCATCT 70
Db 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGATCTCCAGCTGGGACACCTCCAGCGCGGATCTCCAGCCAGGATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyAlaArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGATCTCCAGCTGGGACACCTCCAGCGCGGATCTCCAGCCAGGATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyAlaArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGTACACTCCAGCGCGGATCTCCAGCGCGG-----GCATCT 235
Db 61 ProAlaGlyThrProProGlyAlaArgAlaSerProGlyAlaArgAlaSerProAlaGlnAlaSer 80
QY 236 CCAGCCAGGATCTCCAGCGCGGATCTCCAGCTGTGAGCATCTTCAGGCTCTCA 295
Db 81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaGlnAlaSerLeuSerSerSer 100

```

QY 296 TCCGGCAGTCATCATCCGCGAGGTCAAGCTCGGTGCAAACTCCCAACGAGGTGAC 355
Db 101 SerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
QY 356 CTGTGTAGAGCAACACGATGGGGGCTGTATCCATCCGATCATCTCTCCAGGTGACA 415
Db 121 LeuValAlaArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
QY 416 CCGCAACCGAGGGCCACGAGGAGAGCCGAGTACGAGCTCGCCCAAGTTCACTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgGlySerProGlyThrSerLeuProLysPheThrTyrArg 160
QY 476 GAGGGCCAGAAAGACGATACCGCTCATCGGAGCTGGCTCTCTCATTTAGCCCTGATGAT 535
Db 161 GluGlyGlnGlySerGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValVal 180
QY 536 TCGCTCATCATCTCTTCCAGTTCTTGCAAGGCCACACAGGAGATCAAGTACAGAGACAG 595
Db 181 SerLeuIleIleLeuPheGlnPheThrGlnGlyIleArgTyrGlyGln 200
QY 596 AGGAGAGACCTGTCCCAACGAGCTGTTCGCTGTGACGGGGGTGGAGCTGCAACTGAG 655
Db 201 ArgGlySerCysProLysHisAlaValArgCysArgGlyValAlaAspCysLysLeuLys 220
QY 656 AGTGACGAGCTGGGCTCGTGAGGGTTTGACTGGGACAACTCTGTGTTAAATCTACTCT 715
Db 221 SerThrGlyLeuGlyCysValArgPheAspThrAspLysSerLeuLeuLysIleTyrSer 240
QY 716 GGGTCTTCCCATCATGTGGCTTCCCATCTGTAGACAACTGGAAATGACTCTACTCAGAG 775
Db 241 GlySerSerHisGlnTyrLeuProIleCysSerSerAsnTyrAsnAspSerTyrSerGln 260
QY 776 AAGACCTGGCAAGCTGGGTTTGAGAGTGTCAACGGGACACCGAGGTTGGCCACAG 835
Db 261 LysThrCysGlnGlnLeuGlyPheGlySerAlaHisSerGlnThrGlnValAlaHisArg 280
QY 836 GATTTTGCACAACGCTTCTCAATCTTGAGATACAACTCCACCATCCAGAAAGCTCCAC 895
Db 281 AspPheAlaAsnSerPheSerIleLeuArgGlyThrAsnSerThrIleGlnGlnSerLeuHis 300
QY 896 AGGTGTAAATGCCCTTCCACGGGTATATCTCCCTCAAGTGTCCCACTGGGACTGAG 955
Db 301 ArgSerGlyCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
QY 956 GGCATGACCGGGGGGATGTGTGGAGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAA 1015
Db 321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTyrProTyrGln 340
QY 1016 GTGAGTCTGCACTTCGGACACCACTCATCTGTGAGGACGCTCATTTGACGCCAGTGG 1075
Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTyr 360
QY 1076 GTGCTCACTGCGGCCACTGCTTCTTGTCGACCCGGAGAAAGTCTGGAGGGCTGAAAG 1135
Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgGlyLysValLeuGlnGlyTyrLys 380
QY 1136 GTGTAACCGGGGACACGAACTTCGACGAGTTGGCTGAGCGAGCTCATTTGSCGAGATC 1195
Db 381 ValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIle 400
QY 1196 ATCATCAACAGCAATTAACCGATAGAGAGACGACTATGACATCGCCCTCATGCGGCTG 1255
Db 401 ThrIleAsnSerAsnTyrThrAspGlyGlnAspArgTyrAspIleAlaLeuMetArgLeu 420
QY 1256 TCCAAAGCCCTGACCCCTGTGCTGCTCACTCAACCTGTGCTTCCCATGATGAGACAG 1315
Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
QY 1316 ACCTTTAGCCTCATGAGACCTGTGATCAAGGCTTTGGCAAGACGAGGAGACAGAT 1375
Db 441 ThrPheSerLeuAsnGlnThrCysThrIleThrGlyPheGlyLysThrArgGlnThrAsp 460

```

```

QY 1376 GACAAAGACATCCCCCTTCTCCGGAGGTGACAGTCAATCTCATGCACTTCAAGAAATGC 1435
Db 461 AspLysThrSerProPheLeuArgGlyValGlnValAsnLeuIleAspPheLysCys 480
QY 1436 AATGACTACTTGGTCTATGACAGTTTACTTAACCCAAAGATGATGTGTGCTGGGGACCTT 1495
Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeu 500
QY 1496 CGTGGGGGACAGAGCTCTGCCAGGAGACAGCGGGGGGCTCTGTCTGTGACACAGAAC 1555
Db 501 ArgGlyGlyAlaArgAspSerCysGlnGlyAspSerGlyValProLeuValCysGlnGlnAsn 520
QY 1556 AACCGCTGTACTGTGACAGGTGTCAACGAGCTGGGGCAACAGCTGTGTGGCAAGAAACAA 1615
Db 521 AsnArgThrTyrLeuAlaGlyValThrSerTyrGlyThrGlyCysGlyGlnArgAsnLys 540
QY 1616 CTTGGTGTACACCAAAAGTGACAGAAAGTTCTTCCCTGGATTTTACAGCAAGATGGAG 1672
Db 541 ProGlyValTyrThrLysValThrGluValLeuProTyrIleTyrSerLysMetGln 559

```

Search completed: September 16, 2006, 01:34:33
Job time : 333 secs


```
QY 689 GACAAGTCTGCTAAATCTACTGTGGTCTCCCATGAGTGGCTTCCATCTGTAGC 748
    ::::: |||
Db 59 AlaSerAlaArgLeuMetValPheArgProThrGluThrTrpArgLeuLeuCySer 78
    ::::: |||
QY 749 AGCAACTGGAATGACTCTACTACAGAGAAGACCTGCCAGCGAGCTGGTTTGAGAGTCT 808
    |||
Db 79 SerArgSerAlaAlaArgValAlaGlyLeuSerCysGluGluMetGlyPheLeuArgAla 98
    |||
QY 809 CACCGAGCAACCGAGGTGCCACAGGAT-----TTT 841
    ::::: |||
Db 99 LeuThrHisSerGluLeuSerValAlaArgThrAlaGlyAlaLeuGlyThrSerGlyPhePhe 118
    ::::: |||
QY 842 GCCAAGACCTTCTCAATTTGAGATACACTCCACATCCAGAAAGCTTCCACAGTCT 901
    |||
Db 119 CysValAlaArgGluArgLeuProHISThrGlnArgLeuGluValAlaSerValCys 138
    ::::: |||
QY 902 GAATGGCTTCCCGAGCGGTATATGCTCCCTCGAGTGTCCCATCGCGAGCTGAGGCCATG 961
    ::::: |||
Db 139 AspCysProArgGlyArgPheLeuAlaAlaIleCysGlnAspCysGlyArgArgLeuLeu 158
    ::::: |||
QY 962 ---ACCGGCGGATCGTGGAGAGCGCGCTGGCTCGAGTACCAAGTGGCTTGGCAAGTG 1018
    |||
Db 159 ProValAspArgGlyLeuAlaGlyArgAspThrSerLeuGlyArgTrpProTrpGlnVal 178
    |||
QY 1019 AGTCTGACCTTGGGACACCCACATCTGTGAGGACGCTCATTTGACGCGCAGTGGTG 1078
    |||
Db 179 SerLeuArgGlyArgArgGlyAlaHISLeuGlySerLeuLeuSerGlyAspTrpVal 198
    |||
QY 1079 CTCAGTGGCGGCGGCTGTTCTTCTGTGACCGGAGAGAGTCTGTGAGGCTGGAAGTG 1138
    |||
Db 199 LeuThrAlaAlaHISCys---PheProGluArgAsnArgValLeuSerArgTrpArgVal 217
    |||
QY 1139 TACCGGCGGACCC-----AGCAACCTTGACCAAGTGTCTGTGAGGCGAGCTTCATTGCC 1189
    |||
Db 218 PheAlaGlyAlaValAlaGlnAlaSerProHISGlyLeu-----GlnLeuGlyValGln 235
    |||
QY 1190 GAGATCATCATCAACAGCAATTAC-----ACCGATGAGAGAGACGAC 1231
    |||
Db 236 AlaValAlaValThrGlyGlyLeuProPheArgAspProAsnSerGluGlnAsnSer 255
    |||
QY 1232 TATGACATGGCGCTCATGCGGCTGTGCAAGCCCTGACCCCTGCGGCTCATTCACCTC 1291
    |||
Db 256 AsnArgIleAlaLeuValHISLeuSerSerProLeuProLeuThrGlyTrpIleGlnPro 275
    |||
QY 1292 GCTTGGCTTCCCATGACATGACAGACCTTTAGCCTCATGAGACCTGTGATCAGACGC 1351
    |||
Db 276 ValCysLeuProAlaAlaGlyGlnAlaLeuValAspGlyTrpIleCysThrValThrGly 295
    |||
QY 1352 TTGGGCAAGACCGAGGAGACAGATGACAGACATCCCTTCTCGGAGAGTGCAGGTC 1411
    |||
Db 296 TrpGlyAsnThrGlnTrpGlyGlnGlnAlaGly---ValLeuGlnGlnAlaArgVal 314
    |||
QY 1412 AATTCATCGACTTCAAGAAATGCAATGACTACTGCTGTGACAGTACTTCACTTACCCCA 1471
    |||
Db 315 ProIleIleSerAsnAspValCysAsnGlyAlaAspPheTrpGlyAsnGlnIleLeuPro 334
    |||
QY 1472 AGATGATGTGTGTGGGACCTTCTGTGGGCGAGAGACTCTGCGAGGAGACAGCGGG 1531
    |||
Db 335 LysMetPheCysAlaGlyTrpProGluGlyIleAspAlaCysGlnGlyAspSerGly 354
    |||
QY 1532 GGGCCTTGTGTGTAG-----CAGAACCAACCGCTGTACTTGGCAGGTGTC 1579
    |||
Db 355 GlyProPheValCysGlnAspSerIleSerArgThrProArgTrpArgLeuGlyIle 374
    |||
QY 1580 ACCGCTGGGAGCAGAGCTGTGGCCAGAGAAACAACTGTGTGTATACCAAGTGTGCA 1639
    |||
Db 375 ValSerTrpGlyThrGlyCysAlaLeuAlaGlnLysProGlyValTrpThrLysValSer 394
    |||
QY 1640 GAAGTTCTTCCCTGATTTACAGCAAGATGAGAGC 1675
    |||
Db 395 AspPheArgGluTrpIlePheGlnAlaIleLeuThr 406
    |||
```

```
S33777
hepsin (EC 3.4.21.-) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
C:Accession: S33777; S32013
R:Farley, D., Raymond, F., Nick, H.
Biochim. Biophys. Acta 1173, 350-352, 1993
A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
A:Reference number: S33777; MUID:93305733; PMID:8318546
A:Accession: S33777
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <FAR>
A:Cross-references: UNIPROT:Q05511; UNIPARC:UPI00004ECD9; EMBL:X70900; NID:957928; PIDN:
F:22-44/Domains: hydrolase; liver; serine proteinase; transmembrane protein
F:162-169/Domains: transmembrane #status predicted <TM>
F:187-203,290-358,321-337,348-380/Dissulfide bonds: #status predicted
F:202,256,352/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 2,12e-25 Length: 416
Score: 592.50 Matches: 138
Percent Similarity: 51.8% Conservative: 59
Best Local Similarity: 36.3% Mismatches: 138
Query Match: 17.8% Indels: 45
DB: 1 Gaps: 12

US-10-806-370-11 (1-1748) x S33777 (1-416)
QY 635 GTGGTGAAGTCAAGTGAAGTGAAGTGAAG-----CTGGGCTGCGTGAAGTTGACTGC 688
    ::::: |||
Db 38 IleValThrIleLeuLeuArgSerArgGlnGluProLeuTrpGlnValGlnLeuSerPro 57
    |||
QY 689 GACAAGTCTGCTAAATCTACTGTGGTCTCCCATGAGTGGCTTCCATCTGTAGC 748
    ::::: |||
Db 58 GlyAspSerArgLeuLeuValLeuAspLysThrGluGlyThrTrpArgLeuLeuCySer 77
    |||
QY 749 AGCAACTGGAATGACTCTACTACAGAGAAGACCTGCCAGAGCTGGTTCGAGAGTCT 808
    |||
Db 78 SerArgSerAlaAlaArgValAlaGlyLeuGlyCysGluGluMetGlyPheLeuArgAla 97
    |||
QY 809 CACCGAGCAACCGAGGTGCCACAGGATTT-----GCCACAGC 850
    |||
Db 98 -----LeuAlaHISerGluLeuAspValAlaArgThrAlaGlyAlaAsnGly 112
    |||
QY 851 TTCTCA-----ATCTGAGATACCACTCCACATCCAGGA 886
    |||
Db 113 ThrSerGlyPhePheCysValAspGluGlyLeuProLeuAlaGlnArgLeuLeuAsp 132
    |||
QY 887 AGCTTCCACAGTGTGATGCGCTTCCAGCGGTATATCTCCCTCCAGTGTCCCACTGC 946
    |||
Db 133 ValIleSerValCysAspCysProArgGlyArgPheLeuThrAlaThrCysGlnAspCys 152
    |||
QY 947 GGACTGAGGCGCATG---ACCGGCGGATGCTGGAGAGCGCGCTCGATGACAG 1003
    |||
Db 153 GlyArgArgLysLeuProValAspArgIleValGlyIleGlnAspSerSerLeuGlyArg 172
    |||
QY 1004 TGGCTTGGAGAGTGAATGCTGCACTTGGCAGCACCCACATCTGTGAGGACGAGCTATT 1063
    |||
Db 173 TrpProTrpGlnValSerLeuArgLysArgGlyThrHISLeuGlyGlySerLeuLeu 192
    |||
QY 1064 GACCCCGAGTGGGTCTCATCTGCGCGCCACTCTTCTCGAGACCCGAGGAGAGTCTGT 1123
    |||
Db 193 SerGlyAspTrpValIleLeuThrAlaAlaHISCys---PheProGluArgAsnArgValLeu 211
    |||
QY 1124 GAGGCTGGAAGGTGTACCGCGGC-----ACCAAGCACTGACCAAGTGTGCT 1171
    |||
Db 212 SerArgTrpArgValPheAlaGlyAlaValAlaArgThrSer-----Pro 226
    |||
QY 1172 GAGGCGACCC-----TCCATTGCGGAGATCATCATCAACAGCAATTAC----- 1213
    |||
Db 227 HisAlaValGlnLeuGlyValGlnAlaValIleTrpHISGlyGlyTrpLeuProPheArg 246
    |||
```


Oy 1201 CAACAGCAATTACACCGATGAGAGAGACGACTATGACATGCGCTCATGCGCTGTCCAA 1260
 Db 864 eAaNPohierTyrAsnArgArgLyAspAsnAspIleAlaMetMetHisLeuGluPh 884
 Oy 1261 GCCCTCAACCTCGTCCGCTCATCATCCACCTGCTTCCCTCCCATGATGAGACGACTT 1320
 Db 884 eLyValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGlnValPh 904
 Oy 1321 TAGCCTCATGAGACCTGCTGATCATACAGCTTTGGCCAGACCCAGGAGACAGATGACAA 1380
 Db 904 eProProGlyArgAsnGlySerIleAlaGlyTyrGly---ThrValValTyrGlnGlyTh 923
 Oy 1381 GACATCCCTCTCTCCGCGAGGTGCAGGTCAATTCATGCACTTCAAGAAATCAATGA 1440
 Db 923 rThAlaAsnIleLeuGlnGlnAlaAspValProLeuLeuSerAsnGlnArgGlnG 943
 Oy 1441 CTACTGTGCTATGACAGTTACTTACCCCAAGATGATGTGTCTGTGGGACCTTCGTGG 1500
 Db 943 nGlnMetPro---GluTyrAsnIleThrGluAsnMetIleCysAlaGlyTyrGlnGluG 962
 Oy 1501 GAGCAGAGACTCTGCTCCAGAGAGACAGCGGAGGCTCTTGTCTGTGAGACAGAAACCG 1560
 Db 962 yGlyIleAspSerCysGlnGlyAspSerGlyGlyProLeuMetCysGlnGlnValAsnAr 982
 Oy 1561 CTGCTACTGCGCAGGTGTCTACCACTGCGGCGACAGCTGTGGCCAGAGAAACAACTGG 1620
 Db 982 gTrpPheLeuAlaGlyValThrSerPheGlyTyrIleCysAlaLeuProAsnArgProG 1002
 Oy 1621 TGCTGTACCCAAAGTGAAGAGAGTTCTTCCCTGATTTACAGC 1663
 Db 1002 yValTyrAlaArgValSerArgPheThrGluTyrIleGlnSer 1016

RESULT 4

K0MSPL

plasma kallikrein (BC 3.4.21.34) precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004

C:Accession: A36557

R:Seidah, N.G.; Sawyer, N.; Hamelin, J.; Miron, P.; Beaubien, G.; Brachepa, L.; Rochemot

DNA Cell Biol. 9, 737-748, 1990

A:Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and compariso

A:Reference number: A36557; MUID:91090844; PMID:2264928

A:Accession: A36557

A:Molecule type: mRNA

A:Residues: 1-638 <Seq>

A:Cross-references: UNIPROT:P26262; UNIPARC:UPI0000277BC; GB:M58588; NID:9200358; PIDN:

A:Note: part of this sequence, including the amino ends of both the heavy and light chain

C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w

C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li

A:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology

C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla

F:1-19/Domains: signal sequence; structure predicted <SIG>

F:20-339/Product: plasma kallikrein heavy chain #status experimental <HCH>

F:20-109/Domains: apple repeat <AP1>

F:110-199/Domains: apple repeat <AP2>

F:200-289/Domains: apple repeat <AP3>

F:291-380/Domains: apple repeat <AP4>

F:391-638/Product: plasma kallikrein light chain #status experimental <LCH>

F:391-631/Domains: trypsin homology <TRY>

F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322

F:127,215,308,396,494/Binding site: carbohydrate (asn) (covalent) #status predicted

F:434,483,578/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.: 1,81e-24 Length: 638
 Score: 575.50 Matches: 148
 Percent Similarity: 46.1% Conservative: 65
 Best Local Similarity: 32.0% Mismatches: 129
 Query Match: 17.3% Indels: 120
 DB: 1 Gaps: 19

US-10-806-370-11 (1-1748) x K0MSPL (1-638)

Oy 403 TGCCAGGTCAAGCCACCAAGCCAGGAGGCCAGGTACGAGCTGCCAA 462
 Db 256 CyPheLeuSerThrSerSerArgProSerProPro-----IleProGln 272
 Oy 463 GTTCACTGGCGGAGGCGCAAGACGCTACCGCTATCGGCTGCTCTCTCT--- 519
 Db 273 GluAsnAlaIleSerGlyTyrSerLeuLeuThrCysArgIleThrArgProGluProCys 292
 Oy 520 -----CATTCCTCGGTGGTTCATCATCTCTTCAGTTCGTGGAGG 567
 Db 293 HisSerIleIleTyrSerGlyValAspPheGluGlyGluLeuAsnValThrPhe--- 311
 Oy 568 CCACACAGGATCAGGTACAGAGCA-----GAGGAGAGCTGTCCCAAGC 618
 Db 312 -----ValGlnGlyAlaAspValCysGln-GluThrCysThrIle---Th 325
 Oy 619 TGTTGCTGTGACGGGTGTG-----GACTGCAAGCTGAAGAGTGA 660
 Db 325 rIleArgCysGlnPhePheIleTyrSerLeuLeuProGlnAspCysIle--- 341
 Oy 661 CGAGCTGGCTGC-----GTAGGTTGACTGGACAAGCTCTGCTTAATCTA 711
 Db 342 -GluGluGlyCysIleCysSerLeuArgLeuSerThrAspGlySerProThrArgIleTh 361
 Oy 712 CTGTGGTCTCCCATGAGTGGCTTCCCATCTGTAGCAGCAATGAGTACTCTAC 771
 Db 361 rTyrGly-----MetGlnGlySerSerGlyTyrSe 371
 Oy 772 AGAGAGACTGCGCAGAGCTGGGTTTCAGAGTGTCTACCGAGCAACCGAGTGGCCA 831
 Db 371 rLeuArgLeuCysIleLeuVal-----AspSerProAspCysThrThrIleSile----- 387
 Oy 832 CAGGATTTTGCCAAACGTTCTCAATCTGAGTACACTCCACATCCAGAGAAAGCT 891
 Db 387 ----- 387
 Oy 892 CCACAGTCTGAATGCCCTTCCAGCGGTATATCTCCCTCCAGTGTCCACTGCGGACT 951
 Db 387 ----- 387
 Oy 952 GAGGCGCATACCGGCGGATCTGTGGAGGCGGCTGCGCTCGGATGCAAGTGGCTTG 1011
 Db 388 -----AsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGluTyrProTr 404
 Oy 1012 GCAAGTGAAGTGTGAC-----TTGGCACACCAACCATCTGTGGAGCAGCGCTCAT 1062
 Db 404 rAlnValSerLeuGlnValIleValSerGlnThrHisLeuCysGlyIleSerIleI 424
 Oy 1063 TGAAGCCCACTGGGTGCTCACTGCGGCCCACTGTTCTTGTGTAACCGGAGAAAGTCT 1122
 Db 424 eGlyArgGlnTyrValLeuThrAlaAlaHisCysPhe----- 436
 Oy 1123 GAGAGGC-----TGAAGGTGTACCGCGGACACGCAACTGTGACCA 1164
 Db 437 -AspGlyIleProTyrProAspValTyrPargIleTyrGlyIleLeuSerLeuSerG 456
 Oy 1165 GTTGCCCTGAG-----CGAGCTTCATTTGCGAGATCATATCAACGAATTAAC 1215
 Db 456 uIleThrIleGluThrProSerSerArgIleGlyGluLeuIleIleHisGlnGluTyrLy 476
 Oy 1216 CGATGAGAGAGACGACTAGATGCGCTTCAGGCTGTCCAGGCGCTGACCTGTGC 1275
 Db 476 eValSerGluGlyAsnTyrAspIleAlaLeuIleIleLeuGlnThrProLeuAsnTyrTh 496
 Oy 1276 CGCTCACATCAACCTGCTGCTTCCCTCCCATGATGACAGACACTTTAGCTCAATGAGAC 1335
 Db 496 rGluPheGlnIleProIleCysLeuProSerIleValAspThrAsnThrIleTyrThra 516
 Oy 1336 CTGCTGATCAGAGGCTTGTGCAAGACGAGAGACAGATGACAGATCATCCCTCTCT 1395
 Db 516 nCyTrpValThrIleTyrPheIleTyrThrIleGlu---GlnGlyGluThrGlnAsnIlele 535


```

Oy 1054 CACGCTATTGACGCCGAGGTGGTCTGCTCCGCCCACTGCTTC----- 1099
    |||
    |||
    |||
Db 643 aserleuleseerProaepTlpreValserAlaIahIscyspHeGlnaBpGluThrIi 663
    |||
    |||
    |||
Oy 1100 -TTGCTGACCCGGGAGAAAGTCTCGAGGGCTCGAAGGTGTACCGGGCCACGACCACT 1158
    |||
    |||
    |||
Db 663 ePhelyleTyrSerAepHsiIsthmet-----TPrThralaPheIeuGlyLeu-----Le 679
    |||
    |||
    |||
Oy 1159 GCACACAGTTCCTGAGCAGCTCC-----ATTGCCAGATCATCAT 1200
    |||
    |||
    |||
Db 679 uasPglinseryeasgersAlaserGlyValGlnIuHsielyeulysaAgIleIleIth 699
    |||
    |||
    |||
Oy 1201 CAACAGCAATTACACCGATGAGAGAGACACTATGACATCCGCTTCGCTGCTCCAA 1260
    |||
    |||
    |||
Db 699 rHiePserPheAenAepHethrPheAepTyrAepIleAlaIeuIeuGluLeuGly 719
    |||
    |||
    |||
Oy 1261 GCCCTGACCTCTGCTCCGCTCATACCTGCTGCTCCCTCCCATGATGACAGACTT 1320
    |||
    |||
    |||
Db 719 eProlaIaGluTyrSerThrValaIarPProIleCyseIeuProAepAenThrsvalPh 739
    |||
    |||
    |||
Oy 1321 TAGCCTCAATGAGACCTGCTGATCAGAGCTTGCGAAAGACAGGAGACAGATGACA 1380
    |||
    |||
    |||
Db 739 eProlaIaGlysaIaIeItrPvalThrGlyTyrPoiIahIsthryseGlu-----GlyGlyTh 758
    |||
    |||
    |||
Oy 1381 GACATCCCTCTCTCCGAGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGA 1440
    |||
    |||
    |||
Db 758 rGlyAlaIeulIeuGlnlyrGlyGluIleArGValIleAenGlnThrThryseGluG 778
    |||
    |||
    |||
Oy 1441 CTACTTGGTCTATGACAGTACCTTACCCCAAGATATATGTGTGCTGGGACCTTCGTGG 1500
    |||
    |||
    |||
Db 778 uIeuIeu-----ProGlnGlnIleThrProArGmetwecysValaGlyPheIeuSerG 796
    |||
    |||
    |||
Oy 1501 GGGCAGAGACTCTCTCCGAGGAGACAGCGGGGGGCTCTT---GTCCTGAGACAA 1557
    |||
    |||
    |||
Db 796 yGlyValaPserCyseGlnIyAepSerGlyGlyProIeuSerSerValaGluIysaBpG 816
    |||
    |||
    |||
Oy 1558 CCGCTGATACCTGCGAGGTGTACCAAGCTGGGAGCAGAGCTGTGGCCAGAGAAACAAC 1617
    |||
    |||
    |||
Db 816 yArgIlePheGlnIaGlyValaIserTyrPoiIyGlysaIaGlnArGuaenlyP 836
    |||
    |||
    |||
Oy 1618 TGGTGTGTACCAAGATGACAGAGATTCTTCCCTGATT 1657
    |||
    |||
    |||
Db 836 oGlyValaIyThrArGlyIeProGluValaIarGAspTyrIle 849
    |||
    |||
    |||

```

RESULT 6

KOHUP

Plasma kallikrein (EC 3.4.21.34) precursor - human

N/AIternate names: kininogenin; plasma prekallikrein

C/Species: Homo sapiens (man)

C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Oct-2004

C/Accession: A00921; A37939

R/Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.

Biochemistry 25, 2410-2417, 1986

A/Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four

A/Reference number: A00921; MUID:86243359; PMID:3521732

A/Accession: A00921

A/Molecule type: mRNA

A/Releases: 1-638 <CHU>

A/Cross-references: UNIPROT:P03952; UNIPARC:UPI00000008A; GB:M13143; NID:G190262; PIDN:

R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.

Biochemistry 30, 2050-2056, 1991

A/Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of

A/Reference number: A37939; MUID:91152016; PMID:199666

A/Accession: A37939

A/Molecule type: protein

A/Releases: 20-27/40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80, 103-113, 131-140, 141-

1460-283, 'X', 285-287-291, 'X', 293-295, 314-317, 'X', 319-320, 321-324, 'X', 332-333, 334-339, 'X',

525, 538-551, 562, 'X', 564-567, 573, 'X', 575-576, 578-583, 'X', 585-592-604 <MCW>

A/Cross-references: UNIPARC:UPI0000172842; UNIPARC:UPI0000172843; UNIPARC:UPI0000172844;

849; UNIPARC:UPI0000172844; UNIPARC:UPI0000172845; UNIPARC:UPI0000172846; UNIPARC:UPI0000172847;

10000172852; UNIPARC:UPI0000172853; UNIPARC:UPI0000172854; UNIPARC:UPI0000172855; UNIPAR

C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w

C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 11k are linked by one or more disulfide bonds.
 C/Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r iogen and may also play a role in the renin-angiotensin system by converting prorenin in

C/Genetics: GDB:KLC3
 A/Genes: GDB:KLC3
 A/Cross-references: GDB:127575; OMIM:229000
 A/Map position: 4q35-4q35
 C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; inflan
 F1-19/Domains: signal sequence #status predicted <Sig>
 F120-638/Product: plasma kallikrein #status predicted <Mat>
 F120-390/Domains: plasma kallikrein heavy chain #status predicted <HCH>
 F120-109/Domains: apple repeat <AP1>
 F120-199/Domains: apple repeat <AP2>
 F1200-289/Domains: apple repeat <AP3>
 F1291-380/Domains: apple repeat <AP4>
 F1391-638/Domains: plasma kallikrein light chain #status predicted <LCH>
 F1391-621/Domains: trypsin homology <TRY>
 F121-104, 47-77, 51-57, 111-194, 137-166, 141-147, 201-284, 227-256, 231-237, 292-375, 322-328, 383-
 F1127, 308, 396, 453, 494/Binding site: carbohydrate (asn) (covalent) #status experimental
 F1318-347, 340-345/Disulfide bonds: #status predicted
 F1390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted
 F143, 463, 578/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. NO.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score:	572.50	132	53	101	97
Percent Similarity:	48.3%				
Best Local Similarity:	34.3%				
Query Match:	17.2%				

US-10-806-370-11 (1-1748) x KOHUP (1-638)

```

Oy 596 AGGAGAGCTGTCCCAAGCAGCGCTTTCGCTGTACGGGTGTG----- 640
    |||
    |||
    |||
Db 319 GlnIuThrCyseThryls---MeIleArGcyseGlnPheThrTyrSerIeuLeuPro 337
    |||
    |||
    |||
Oy 641 ---GACTGCAAGCTGAGAGTGAAGAGCTGGCGCTG---GTGAGTTTGAAGTGAAG 694
    |||
    |||
    |||
Db 338 GluAepCyse---LysGluGlnIuCyseLysCysePheLeuArGlyLeuSerMetAepI 355
    |||
    |||
    |||
Oy 695 TCTCTGCTTAATATC-----TACTTGGGTCTCTCCATCAAGTGGCTTCCATCTGT 745
    |||
    |||
    |||
Db 356 SerProthArGlyleAlaTyrGlyThrGlnIySerSerGlyTyrSerLeuArGlyCyse 375
    |||
    |||
    |||
Oy 746 AGCAGCACTGGAATGACTCTTACTCAGAGAAGACCTGCCAGCAGCTGGGTTTGAAGT 805
    |||
    |||
    |||
Db 376 AenThrGlyAepAenSerValCyseThrThrylsThr----- 387
    |||
    |||
    |||
Oy 806 GCTACCGAGCAACGAGGTGGCCACAGGAGATTTCGCAACAGCTTCTCAATCTTGAGA 865
    |||
    |||
    |||
Db 387 ----- 387
    |||
    |||
    |||
Oy 866 TACAATCCACATCAGAGAAAGCTCCACAGGTGTGATATGCCCTCCACAGGTATATC 925
    |||
    |||
    |||
Db 387 ----- 387
    |||
    |||
    |||
Oy 926 TCCTCCAGTGTTCCTCACTGGGACGTAGAGGCCATGACCGGGCGGATCTGGAGGGGG 985
    |||
    |||
    |||
Db 388 ----- 388
    |||
    |||
    |||
Oy 986 CTGGCTCGATAGCAGTGGCTTGGCAAGTGTGCTGCACTTCGGCAGCAGC----- 1039
    |||
    |||
    |||
Db 396 AenSerSerTyrGlyIuTyrProIlePheIuValserIeuGlnIuLysIeuThralaGln 415
    |||
    |||
    |||
Oy 1040 ---CACATCTGTGAGGACAGCTCATTTAGAGGCCAGGTGGTGTCTACTGGCCCACTGC 1096
    |||
    |||
    |||
Db 416 ArgIstIeuCyseGlyGlySerIeuIleGlyHsiGlnIuThrValleuThralaIahIscys 435
    |||
    |||
    |||
Oy 1097 TTCTTCGTGACCCGGGAGAAAGTCTCGAAGGC-----TGGAAGTG 1138
    |||
    |||
    |||

```



```

QY      826 TGCCACAGGATTTGGCAACAGCTTCTCAATCTGAGATACAACCTCCACCATCCAGCA 885
D8      387 e-----
QY      886 AAGCTCCACAGGCTGTAATGCCCTTCCACGGATATATCTCCCTCCAGTGTCCCACTG 945
D8      387 -----
QY      946 CGAGTAGAGGCGCATGACCGGCGGATGCTGGAGAGGCGCTGGCTGCGATGACCAATG 1005
D8      388 -----
QY      1006 GCCTTGGCAAGTAGTGTGAC-----TTCGGCACCAACCAATCTGTGAGAGC 1056
D8      402 pPrtGlnValSerLeuGlnValLysLeuValSerGlnAanHisMetCysGlyGlySe 422
QY      1057 GCTATGATGACCCCGAGGCGGCTCACTGCGGCCCACTGCTTCTTCGACCCGGAGAA 1116
D8      422 rTleIleGlyValGlnTrpIleuThrAlaHisCysPhe----- 436
QY      1117 GATCTGAGAGGC-----TGAAGGTGATACCGGCGCACCAACACT 1158
D8      437 -----AspGlyIleProIytrProAspValTTPArgIleTyrGlyGlyIleLeuAenLe 454
QY      1159 GCACCAAGTGCCTAGACA-----GCTTCATTCGAGATCATCATCAACAGCAA 1209
D8      454 uSerGluIleThrAsnIythrProPheSerSerIleGlyGluIleHisGlnIly 474
QY      1210 TTACACGATGATGAGAGACGATGACATGCGCCTCAATGCGGCTGCCAAGCCCTGAC 1259
D8      474 sTyrLysMetSerGlnIySerTyrAspIleAlaLeuIleLysLeuGlnThrProLeuAs 494
QY      1270 CTTGTCGCTCATCATCCACCTGCTGCTCCCATGATGATGACAGACACTTACCTCAA 1329
D8      494 nTyrThrGluPheGlnIySerProIleCysLeuProSerIyAlaSerHisThrAenThrIleTy 514
QY      1330 TGAGACTGCTGTGATCACAGGCTTTGGACAGACAGAGACAGACAGATCCACC 1389
D8      514 rThraAsnCysTrpValThrGlyTyrGlyTyrThrLysGlnIyLysGlnIyLysGlnIy 533
QY      1390 CTTCTCCGGAGAGTGACGAGCATCTCATGACATTCATGACATTCAGAAAGGC-----AA 1437
D8      533 nIleLeuGlnIyAspIleThrIleProLeuValProAenGluGlnIyLysIyTyrAr 553
QY      1438 TGACTACTGTGTATGATCAAGTTACTTACCCCAAGATGATGCTGTGGAGACTTGC 1497
D8      553 gAspTyrValIle-----ThrLysGlnMetIleCysAlaGlyTyrIyLysG 568
QY      1498 TGGGGGACAGAGATCTCTGCGACAGAGACAGCGGGGGGCTTGTCTGTGACAGAA 1557
D8      568 uGlyGlyIleAspIleCysLysGlyAspSerGlyIyProLeuValCysLysHisSerG 588
QY      1558 CCGGTGATGATCTGCGAGGTGTCACGAGTGGGACAGCGCTGGCGCCAGAAACCAAC 1617
D8      588 yArgTyrGlnLeuValGlyIleHisSerTyrGlyGlnGlyCysAlaArgLysGlnIyPr 608
QY      1618 TGGGTGTATACCAAGATGACAGAGAGTCTTCCCTGATTTACAGCAAGATGAGAGC 1675
D8      608 oGlyValIyThrIySerValAlaGluTyrIleAspTrpIleLeuGlnIySerIleGlnSer 627

RESULT 8
A53663
Enteropeptidase (EC 3.4.21.9) precursor [validated] - pig
N:Alternate names: enteropeptidase
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 28-Apr-2003
C/Accession: A53663
R:Matsumura, M.; Ichinose, M.; Yahagi, N.; Kakei, N.; Tsukada, S.; Miki, K.; Kurokawa,
J. Biol. Chem. 269, 19976-19982, 1994
A:Title: Structural characterization of porcine enteropeptidase.
A:Reference number: A53663; MUID:94327546; PMID:8051061
A/Accession: A53663

```

```

A:Molecule type: mRNA
A:Residues: 1-1034 <MAT>
A:Cross-references: UNIPARC:UPI0000172B0C; GB:D30799; NID:G505122; PIDN:BA06459.1; PID:5
A:Note: parts of this sequence, including the amino ends of three chains isolated from th
A:Comment: The mechanism of association with the membrane of the intestinal brush border
C:ated below) or with amino-terminal myristoylation of the heavy chain.
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involve
C:Function:
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
A:Pathway: intestinal digestive hydrolase cascade
C:Superfamily: enteropeptidase; Ctr/Cts repeat homology; LDL receptor ligand-binding repe
C:Keywords: glycoprotein; hydrolase; serine proteinase; trypsinogen; zymogen
F:12-38/Domain: transmembrane #status predicted <TM>
F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>
F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
F:119-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:357-519/Domain: MAM homology <MAM>
F:541-646/Domain: Ctr/Cts repeat homology <CTR>
F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRCT
F:800-1034/Product: enteropeptidase light chain #status predicted <LC>
F:116-147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,902,964
F:781-911,925-841,925-992,996-971,982-1010/Disulfide bonds: #status predicted
F:840,891,986/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 2,136-23 Length: 1034
Score: 556.00 Matches: 143
Percent Similarity: 46.7% Conservative: 56
Best Local Similarity: 33.6% Mismatches: 164
Query Match: 16.7% Indels: 63
DB: Gaps: 15

US-10-806-370-11 (1-1748) x A53663 (1-1034)
QY      475 GAGAGGCGCCAGAGCA-----GTAACCGCTCATCGGCTGCTGCTCTCTCATTTGCCCT 528
D8      640 GlyIyLysPheLysAlaAenPheThrThrGlyTyrHisLeuGlyIleProGluProCys--- 658
QY      529 GGTGTTTGGCTCATCATCTCTTCCAGTTCTGCGACGGCCACAGAGATCATAGTCA 588
D8      659 -----LysGluAAspAenPheGln 664
QY      589 GAGACAGAGGAGAGAGTGTCCCAAGACAGCGTTCGCTGACCGGGGTGTGACAGTCCA 648
D8      665 CysGluAenGlyGlu-CysValLeuLeuValAenLeuCysAspIyLysHisCysGly 684
QY      649 GCTGAGAGTGACAGCTGGGCTGCGTGGAGTTTGAATGGGACAGAGTCTGCTTAAAT 708
D8      684 sAspGlySerAspIyLysAlaHisCysValArgPhe----- 695
QY      709 CTACTCGAGTCCCTCCCATGAG-----TGGCT 735
D8      696 -LeuAenGlyThrAlaAenAenSerGlyLeuValGlnPheArgIleGlnSerIleTrpHis 715
QY      736 TCCCATGTGATGACAGCACTGGAATGACTCTCACTCAGAGAAAGCTGCGACAGTGGG 795
D8      715 sThrAlaCysAlaGlnAenTrpThrGlnThrHisSerAspAspAlaCysGlnLeuLeuG 735
QY      796 TTTGAGAGTGTCTACCGGACCAACCGAGGTGGCCACAGGAGATTGGCAACAGCTTCTC 855
D8      735 yLeuGlyThrGlyAAsn-----SerSerMetProPhePheSerSerGlyG 750
QY      856 A-----ATCTGAGATATCACTCCACATCCAGGAAGCTC-----CACAGTTC 900
D8      750 yGlyProPheValIyLeuAenThrAlaProAenGlySerLeuIleLeuThrAlaSerG 770
QY      901 TGAATGCCCTTCCACAGGAGTATATCTCCCTCCAGTGTCCAC-----TGCGGA----- 949
D8      770 uGlnCysPheGlnAspSerLeuIleLeuLeuGlnCysAanHisIyLysSerCysGlyLys 790

```



```

Db      764 rLeuileuThrProserGlnGlnCysLeuGluAspSerLeuileuLeuGlnCysAs 784
Qy      940 CCAC-----TGCGAATGAGGCGCATGACC-----GCGCGATCGTGGAGG 981
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      784 nTylYserSerCysGlyYsYsLeuValThrGlnGluValSerProYsIleValGlyG 804
Qy      982 GCGCGTGGCTCGATAGCAAGTGGCTTGGCAAGTGATGATGCACTTGGCACCACCA 1041
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      804 YserAspSerArgGlnGluAlaTrpCotrpValAlaLeuYtrPheAspArgGlnG 824
Qy      1042 CATCTGGGAGGAGCAAGCTTATGAGCGCCAGTGGGTGTCTCACTCCGCCCTGCTT 1101
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      824 nValCysGlyAlaSerLeuValSerArgAspTrpLeuValSerAlaAlaHisCysValTy 844
Qy      1102 CGTGAACCGGAGAGAGGCTCTGAGAGGCTGGAAGTGATGACCGGAGC-----ACCAG 1152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      844 r---GlyArgAsnMetGluProserIystrIyAlaValLeuGlyLeuHisMetAlaSe 863
Qy      1153 CAACCTGACCAAGTTGGCT--GAGGACGCTTCATGCGCCAGATCATCATCAACAGCA 1209
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      863 rAsnLeuThrSerProGlnIleGlnThrArgLeuIleAspIleValIleAsnProH 883
Qy      1210 TTAACCGATAGAGAGACGATGACATGACATGCGCTCATGCGGCTGTCCAGCCCTGAC 1269
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      883 rTyrAsnIyAsArgIyAsnAsnAspIleAlaMetMetHisLeuGluMetIyValAs 903
Qy      1270 CCTCTCGGCTCATCCACCTGCTGCTGCTCCCATGATGACAGACCTTACCTTAA 1329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      903 nTylThrAspTrpIyIleGlnProIleCysLeuProGlnGluAsnGlnValPheProProG 923
Qy      1330 TGAACCTGCTGATACAGGCTTGGCAGAGACCAAGGAGACAGATGACAAAGATCCCC 1389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      923 YArgIleCysSerIleAlaGlyTrpGly---AlaLeuIleTyrGlnIySerThrAlaAs 942
Qy      1390 CTTCCTCCGGAGAGGACAGTCATCTCATGACTTCAAGAAATGCAATGACTTCTGCT 1449
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      942 pValLeuGlnGlnIleAlaAspValProLeuLeuSerAsnIlyySerGlnGlnImePr 962
Qy      1450 CTATGACAGTTTACCTTACCCCAAGAGATGATGTGTCTGGGAGCCTTCCTGGGGGACAG 1509
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      962 o---GlnTyrAsnIleThrGluAsnMetValCysAlaGlyTyrGluAlaGlyGlyValAs 981
Qy      1510 CTCTGCGAGGAGAGACAGCGGGGGGCTTGTCTGTGTGAGAGAAACAACCGCTGTA 1569
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      981 pSerCysGlnIyAspSerGlyGlyProLeuMetCysGlnGlnIyAsnAsnArgTrpLeu 1001
Qy      1570 GGCAGGCTGTCACAGCTGGGGGACAGGCTGGGCCAGAGAAACAACCGTGTGTGTA 1629
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1001 uAlaGlyValThrSerPheGlyTyrGlnCysAlaLeuProAsnArgProGlyValTyAl 1021
Qy      1630 CAAAGTACAGAAAGTTCTTCCCTGGATTTACAGC 1663
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1021 aArgValProArgPheThrGluTrpIleGlnSer 1032

```

```

F:447-482/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:648-682/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:684-720/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:723-757/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:865-1097/Domain: trypsin homology <TRY>

Alignment Scores:
Pred. No.:      5,18e-23      Length:      1113
Score:          549.00      Matches:      137
Percent Similarity: 51.6%      Conservative: 60
Best Local Similarity: 35.9%      Mismatches: 135
Query Match:    16.5%      Indels:      50
                   2      Gaps:      18

US-10-806-370-11 (1-1748) x JE0315 (1-1113)
Qy      605 TGTCCCAAGACACGCTGTCCCTGTGAGAGGGGTGGTGGACTGGCAAGCTGAAGTGAAG 664
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      735 CysValProArgAspLeuTrpCysAspGlyTrpValAspCysSerAspSerAspGlu 754
Qy      665 CTGGGCTGCGTGAAGTTTGACTGGAGC-----AAGTCTGTCTTAAATCTAC-- 712
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      755 TrpGlyCysValThrLeuSerIyAsnGlyAsnSerSerLeuLeuThrValHisIy 774
Qy      713 TCTGGGTCTCCCATCAGTGGCTTCCATCTGTAGCAAGCACTGGAATGACTCTTACT 772
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      775 SerAlaIyAsnGlnHis-----ValCysAlaAspGlyTrpArgGluThrLeuSer 791
Qy      773 GAGAAAGCTGCCAGACGCTGGGTGGAGAGTGTGACAGAGTGTCAACGGAGACGAGTTGCC 832
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      792 GlnLeuAlaCysIyGlnMetGlyLeuGlyGluProSerValThrIyLeuIleProGly 811
Qy      833 AGGAGT-----TTGCAACAGCTTCTCATCTTGAATGATAGACAA 871
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      812 GlnIyGlnGlnIyGlnIyTrpLeuArgLeuTyrProAsnTrpGluAsnLeu-----AsnGly 829
Qy      872 TCCACCATCCAGAGAACCTC--CACAGCTGTAATGCCCTTCCACGCGTATATCTCC 928
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      830 SerThrLeuGlnGlnIyLeuLeuValTyArgHisSerCysProSerArgSerGluIleSer 849
Qy      929 CTCAGGTGCC-----CACTGGGACGTGAG--GCCATGACCGGGCGGATC 973
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      850 LeuLeuCysSerIyGlnAspCysGlyArgArgProAlaAlaArgMetAsnIyAsyIle 869
Qy      974 GTGGAGGGGCGCGCTGCGCTGATAGCAAGTGGCTTGGCAAGTGAAGTGTGCACTTGGC 1033
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      870 LeuGlyGlyArgHisSerArgProGlyArgTrpProIyGlnCysSerLeuGlnSerGlu 889
Qy      1034 AACAC--CACATCTGTGAGGACGCTGTAAGCCGCCAGGTGGGTCTCACTGGCGCC 1090
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      890 ProSerGlnHisIleCysGlyCysValLeuIleAlaIyAsyIyTrpValIleValAla 909
Qy      1091 CACTGCTTCTTCGTGACCCCGGAGAAAGTCTGTGAGGGC-----TGG 1132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      910 HisCysPhe-----GlnIyArgGluAspValAlaTrp 921
Qy      1133 AAGGTGACGGGACACAGCACTGACACAG-----TTGCTGAGGACGCTCC 1183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      922 IyValValIleGlyIleAsnAsnLeuAspHisProSerGlyPheMetGlnThrArgPhe 941
Qy      1184 ATTGCGCAATCATCATCAACAGCAATTTACACCGATAGAGAGCACTGATGACATGCGC 1243
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      942 ValIyThrIleLeuLeuHisProArgTySerArgAlaValIyAspTrpIleSer 961
Qy      1244 CTCAATGGGCTGTGTCAGCCCGCTGACCTGTGCTGACATTCACCGCTGTGCTCC 1303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      962 ValValGluLeuSerAspAspIleAsnGlnThrSerTyValArgProValCysLeuPro 981
Qy      1304 ATGATGAGACAGACCTTTAGCTTCATAGAGACCTGTGATACAGGCTTGGCAAGAC 1363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      982 SerProGlnGlnIyTrpLeuGluTrpAspThrTyCysArgIyIleThrGlyTrpGlyHisMet 1001
Qy      1364 AGGAGACAGATGACAAAGACATCCCTTC--CTCCGGAGGTGACAGTCAATGTATC 1420

```

```

Db 1002 -----GlyAlaMetProPheLeuGlnGluValArgIle11017
Oy 1421 GACTTCAAGAAATGCAATGACTTGGTCTATGAC--AGTTACTTACCCCAAGATG 1477
Oy 1018 ProLeuGlnGlnGlnSerTyr-----PheAspMetLysThrIleThrAsnArgMet 1035
Oy 1478 ATGTGTCTGGGGGCTTTCGTCGGGAGAGACTTCCTCCAGGAGAGACAGCGGGGGCTT 1537
Oy 1036 IleCysAlaGlyTyrGlnSerGlyThrValAspSerCysMetGlyAspSerGlyPro 1055
Oy 1538 CTTGTCTGTGAGCG--AACACCGCTGTACTGTGAGCTGCACAGCTGAGGAGACA 1594
Oy 1056 LeuValCysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1075
Oy 1595 GCGTGTGCGCAGAGA--AACAACTGTGTGTATACCAAGAGAGAGAGATTCTTCCC 1651
Oy 1076 ValCysPheSerLysValLeuGlyProGlyValTyrSerAsnValSerTyrPheValGly 1095
Oy 1652 TGGATT 1657
Oy 1096 TriPle 1097

RESULT 11
JC5759
brain-specific serine proteinase (EC 3.4.21.-) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 31-Dec-2004
C/Accession: J05759
R/Yamamura, Y.; Yamashiro, K.; Tsunoka, N.; Nakazato, H.; Tsujimura, A.; Yamaguchi, N.
Biochem. Biophys. Res. Commun. 239, 386-392, 1997
A/Title: Molecular cloning of a novel brain-specific serine protease with a kringle-like
A/Reference number: J05759; MUID:98008848; PMID:9344839
A/Accession: J05759
A/Molecule type: mRNA
A/Residues: 1-761 <YAM>
A/Cross-reference: UNIPARC:UP10000175C77; DDBJ:ID89871
A/Experimental source: brain
C/Superfamily: plasma hyaluronan-binding protein; scavenger receptor cysteine-rich domain
C/Keywords: glycoprotein; hydrolase; serine proteinase
F/85-157/Domain: kringle-like #status predicted <KRI>
F/163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F/166-266,273-372,386-486/Domain: scavenger receptor cysteine-rich #status predicted <SR>
F/513-516/Domain: furin binding #status predicted <TRY>
F/517-755/Domain: tryptsin homology
F/93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/562,612,711/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 1.64e-22 Length: 761
Score: 540.00 Matches: 194
Percent Similarity: 34.8% Conservative: 73
Best Local Similarity: 25.3% Mismatches: 240
Query Match: 16.2% Indels: 262
DB: 2 Gaps: 30

US-10-806-370-11 (1-1748) x J05759 (1-761)
Oy 59 GCTGAGCATCTCCAGCCAGCATCTCCAGCTGGAGACCT-----CCAGCCGG 109
Oy 19 AlaArgAlaPheProValSerArgSerProLeuHisArgProHisProSerProArg 38
Oy 110 GCATCTCCAGCCAGCATCTCCAGCCAGCATCTCCAGCTGGAGACCT-----160
Oy 39 SerGlnHisAlaHisTyrLeuProSerSerArgArgProArgThrProArgPhePro 58
Oy 161 -----CCGGCGCCGCGCATCTCCAGCCAGCATCTCCAGCTGGAGACCT 205
Oy 59 LeuProLeuArgIleProAlaAlaGlnArgProGlnValLeuSerThrGlyHisThrPro 78
Oy 206 -----CCAGCGCCGCGCATCTCCAGCCAGCCAGCATCTCCAGCTGGAGACCT 226
Oy 79 ProThrIleProArgArgCysGlyAlaGlyGlnSerTyrGlyAsnAlaThrAsnLeuGly 98

```

```

Oy 227 -----CGGGCATCTCCAGCCAG 244
Oy 99 ValProCysLeuHisTyrPheArgIleValProProPheLeuGlnLysSerProProAlaSer 118
Oy 245 GCATCTCCAGCCAGCATCTCCAGCTGGAGACCTTTCAGAGTTCATCCGCGAG 304
Oy 119 ThrAlaGlnLeuArgGlyGlnProHisAsnPheCysArgSerProArgGlySerGlyArg 138
Oy 304 -----304
Oy 139 ProTyrPheTyrArgAsnAlaGlnGlyValValAspTyrGlyTyrCysAspCysGly 158
Oy 305 -----TCATCATCCCGCCAGCTGAGCTCG 328
Oy 159 GlnGlyProAlaLeuProValIleArgLeuValGlyLysAsnSerGlyHisGlnGly---177
Oy 329 GTGACAACTCCCAACCAAGAGTGTACTTGTAGACAAACACCAAGTGGGGCTGTACC 388
Oy 178 -----ArgValGlnLeuTyrHisAla-GlyGlnTyrGlyThrIleCys 191
Oy 389 ATCCGATCAT-----398
Oy 191 AspAspGlnTyrPheAsnAlaAspAlaAspValIleCysArgGlnLeuGlyLeuSerGly 211
Oy 399 -CTCCCTCCAGCTCAGCAGCAGCA-----CCAGGCGCAGCGGAGAGAC 442
Oy 211 YIleAlaLysAlaTyrPheHisGlnAlaHisPheGlyGlnGlySerIleProIleLeuLys 231
Oy 443 CCAGGTACGAGCGCTG-----CCCAAGTTCACCTGGCG 474
Oy 231 PGIValAlaArgCysThrGlyValGlnGlnMetSerIleGlnGlnCysProLysSerSerTyrGly 251
Oy 475 GGAG-----GGCAGAGCAGCTACCGCTGAGCTGGCTCTCTCATTTGC 525
Oy 251 YGlnHisAsnCysGlyHisLysGlnAspAlaGlyValSerCysValProLeuThrAspGly 271
Oy 526 CTTGGTGTTCCTGCTATCATCTCTTCACCTTTCGACAGGCGCAGCAGGATCAGGTA 585
Oy 271 Y---ValIleArgLeuLysGlnGlyLysSerThrHisGlnGlnArgLeuGlnValTyrTyr 290
Oy 586 CAAGGAGCAGAGGAGAGC--TGTCACAGCAGCTGTTGCTGTGACGCGGCTGTGCA 642
Oy 290 KlyValGlnTyrGlnTyrValCysAspArgGlyTyrThrGlnMetAsnThrTyrValAl 310
Oy 643 CTGCAACTGAAAGTGCAGCAGCTGGC-----670
Oy 310 AcYArgLeuLeuGlyPheLysTyrGlyLysGlnSerSerValAsnHisPheAspGlySe 330
Oy 670 -----670
Oy 330 PAsnArgProIleTyrLeuAspAspValSerCysSerGlyLysGlnValSerPheIleGly 350
Oy 671 -TGGGTAGGTTTACTCG-----GACAAGTCTTG--700
Oy 350 nCysSerArgArgGlnTyrGlyArgHisAspCysSerHisArgGlnAspValGlyLeuTh 370
Oy 701 -----CTTAATCTACTC 714
Oy 370 rCysTyrProAspSerAspGlyHisArgLeuSerProGlyPheProIleArgLeuValAs 390
Oy 715 TGGGTCTCCCAT-----CAGTGGCTCCCAT 741
Oy 390 PGIValAsnLysLysGlnGlyArgValGlnValPheValAsnGlyGlnTyrGlyThrIle 410
Oy 742 CTGTAGCAGCAACTGGAATGACTCTACTCAGAGAAACCTGCAGCAGCTGAGTTTGA 801
Oy 410 eCysAspAspGlyTyrThrAspLysHisAlaAlaValIleCysArgGlnLeuGlyTyrLys 430
Oy 802 GAGTGTCAACCGGCAACCAAGGTTGCCACAGGAGATTTGCCAACAGCTTCAATTT 861
Oy 430 sGlyProAlaArgAlaArgThrMetAlaTyr-----PheGlyGlnGlyLysGlyProIle 448

```

```
QY 862 GAGATACAACTCCAC-----ATCCAGCA 885
      :
      :
      :
Db 448 AHEMetAspSerValLysCysThrGlyAsnGluLysAlaLeuAlaAspCysValLysG 468
      :
      :
      :
QY 886 AAGCTCCACAGGCTGTAATGCCCTCCAGCGGTATATCTCCCTCCAGT----- 937
      :
      :
      :
Db 468 nAspIleGlyArgHisAsnCySerHisSerGluAspAlaGlyAlaIleCysAspTyrLe 488
      :
      :
      :
QY 938 -----TCCCATGCTGGAGTCAAG 954
      :
      :
      :
Db 488 uGluLysAlaLaserSerSerGlyAsnLysGluMetLeuSerSerGlyCysGlyLeuAr 508
      :
      :
      :
QY 955 GGCCATG-----ACCGGCGGATCTGTGGAGAGCGCGCTGCTCGATGACCAAGTG 1005
      :
      :
      :
Db 508 gLeuLeuHisArgArgGlnLysArgIleIleGlyGlyAsnAsnSerLeuArgGlyAlaTr 528
      :
      :
      :
QY 1006 GCCTTGGCAAGTGAAGTCTGCACTTCGGACCAACCCAC-----ATCTGTGG 1050
      :
      :
      :
Db 528 pProTyrPdlHisSerLeuArgLeuArgSerAlaHisIleGlyAspGlyArgLeuLeuCyse 548
      :
      :
      :
QY 1051 AGGCAAGCTCATTTGACCGCCAGATGCTGCTGACCTGCGCCGCTCTTCTGTCGACCCG 1110
      :
      :
      :
Db 548 yAlaThrLeuLeuSerSerCysTyrValLeuThrAlaAlaHisCysPhe----- 564
      :
      :
      :
QY 1111 GGAGAAAGTCTGAGAGGCTGAAAGGTGTACCGCGGACCAAGCACTGCAC----- 1162
      :
      :
      :
Db 565 -----LysArgTyrGlyAsnAsnSerArgSerTyrAlaValAr 577
      :
      :
      :
QY 1163 -----CAATGCTTGAAGCAGCTTCATTCGCGA 1191
      :
      :
      :
Db 577 gValGlyAspTyrHisThrLeuValProGluGluPheGluGlnGluIleGlyValGlnG 597
      :
      :
      :
QY 1192 GATCATCTCAACAGCACTTACACCGATGAGAGAGAGACGATGACATGCGCCCTCATCG 1251
      :
      :
      :
Db 597 nIleValIleHisArgAsnTyrArgProAspArgSerAspTyrAspIleAlaLeuValAr 617
      :
      :
      :
QY 1252 GCTGTCCAAAGCC-----CTGACCTGTCTGCTGACATCCACCTGCTTGGCT 1299
      :
      :
      :
Db 617 gLeuGlnGlyProGluGlnGlnCysAlaArgLeuSerThrHisValLeuProAlaCysLe 637
      :
      :
      :
QY 1300 CCCCATG-----CATGACAGACCTTAAAGCTCAATGACAGCTGTGATGATCAAGG 1350
      :
      :
      :
Db 637 uProLeuTyrArgGluArgProGlnLysThrAlaSerAsn-----CysHisIleThrG 655
      :
      :
      :
QY 1351 CTTTGGCAAGACGAGAGAGACAGATGACAGACATCCCTTCCCTCCGGAGGTGACAGT 1410
      :
      :
      :
Db 655 yTyrGlyAspHisArgHisArgAlaTyrSerArgThr-----LeuGlnGlnAlaAlaVa 672
      :
      :
      :
QY 1411 CAATCTCATGACTTCAAGAAATGCAATGACTACTGTGCTATGACAGTTACCTTACCC 1470
      :
      :
      :
Db 672 lProLeuLeuProLysArgPheCysLysGluArg-----TyrLysGlyLeuPheThrG 690
      :
      :
      :
QY 1471 AAGATGATGTGTCTGTTGAGACCTTCGTGGGGGCGA-----GATCCTGCTCCAGGAGA 1524
      :
      :
      :
Db 690 yArgMetLeuCySerAlaGlyAsnLeuGlnLysAspAsnArgValAspSerCysGlnGlyAs 710
      :
      :
      :
QY 1525 CAGGAGGGGGCTTGTCTGTGAGCAG---AACACCGCTGTACTGACGAGGTGTTCAC 1581
      :
      :
      :
Db 710 pSerGlyGlyProLeuMetCysGluLysProAspGluSerTyrValValTyrGlyValTh 730
      :
      :
      :
QY 1582 CAGCTGGGGACAGAGCTGTGGCCAGAGAAACAACTGTGTGTACACCAAGTGACAGA 1641
      :
      :
      :
Db 730 rSerTyrPdlTyrArgCysGlyValAspThrProGlyValTyrThrArgValProAl 750
      :
      :
      :
QY 1642 AGTTCTTCCCTGATTTTACAGC 1663
      :
      :
      :
Db 750 aPheValProTyrPdlLeuSer 757

RESULT 12
A47547
serine proteinase stubble-stubloid (BC 3.4.21.-) - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
```

```
C/Accession: A47547
R/Appel, L.F.; Procut, M.; Abu-Shumays, R.; Hammond, A.; Garbe, J.C.; Friele, D.; Friele,
Proc. Natl. Acad. Sci. U.S.A. 90, 4937-4941, 1993
A/Title: The Drosophila stubble-stubloid gene encodes an apparent transmembrane serine I
A/Reference number: A47547; MUID:93281671; EUID:7685111
A/Accession: A47547
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-786 <APP>
A/Cross-references: UNIPROT:Q05319; UNIPARC:UP1000013615A; GB:L11451; MID:G158511; PIDN:
C/Genetic:
A/Gene: Sb-rbd
A/Cross-references: FlyBase:FBgn000319
C/Superfamily: serine proteinase stubble-stubloid; trypsin homology
C/Keywords: hydrolase; serine proteinase; transmembrane protein
F:61-77/Domain: transmembrane #status predicted <TMN>
F:543-781/Domain: trypsin homology <TRY>

Alignment Scores:
Pred. No.: 1,74e-22 Length: 786
Score: 539.50 Matches: 181
Percent Similarity: 46.7% Conservative: 92
Best Local Similarity: 31.0% Mismatches: 224
Query Match: 16.2% Indels: 88
DB: 1 Gaps: 23

US-10-806-370-11 (1-1748) x A47547 (1-786)

QY 8 ACCATGAGAGAGGACACCCAGGGAATGATCTCCAGCAAGAACACCTTGACGTGAGCA 67
      :
      :
      :
Db 254 ThrThrAlaSerAspLeuHisGlySerAlaSerHisProSerSerSerSerSerSer 273
      :
      :
      :
QY 68 TCTTCACCCAGGAGATCTCCAGCTGAGAGACCTTCAGCGGAGATCTCCAGCCAGCA 127
      :
      :
      :
Db 274 SerSerSerSerProAsn-serIleTyrHisThrSerThrGlnGlnGlnGlnGlnH 293
      :
      :
      :
QY 128 TCTTCACCCAGGAGATCTCCAGCTGAGAG---ACCTCCGGAGCGGAGC---ATCTCCAGCC 181
      :
      :
      :
Db 293 sGlnGlnAlaAsnGlnAla-AsnHisIleTyrGlnMetThrThrGluProSerPheIleThrLysP 313
      :
      :
      :
QY 182 CAGGCACTCTCCAGTGTACACCTTCAGCGGAGCACT---CGAGGCGGGGATCTCCA 238
      :
      :
      :
Db 313 roArgProThrGlyTyr-ThrLysProGlyTyrLeuAlaAsnLeuProMetProAlaArgPro 332
      :
      :
      :
QY 239 GCCCAGGACCTTCAGCGCGGAGCATCTCCGAGCTGTGACATCATCTTCCAGGTCTCATCC 298
      :
      :
      :
Db 333 SerLysProSerLysProThrLysLysProIleVal-----TyrAspArgSerProPro 350
      :
      :
      :
QY 299 GGCAGTCAATCATCCGACAGGTCAAGCTCGGTGACAACTCCCAACCAAGATACCTT 358
      :
      :
      :
Db 351 ProProSerSerValProProSerThrSerThrSerThrSerThrSerThrSerLeuIleTyr 370
      :
      :
      :
QY 359 GTTAGACAAACACAGAGGGGGCTGTACCATCCGATCATCTCTGCGCAGGTACACCA 418
      :
      :
      :
Db 371 -----ProAlaGlnThrHisPro-Pro----- 377
      :
      :
      :
QY 419 GCAACAGGGGACACAGAGAGAGCCAGTACAGAGCTTCCCAAGTTCACTGCGGAGAG 478
      :
      :
      :
Db 378 -GlnProHisArgProThrArgProGlnLeu-----SerProGly----- 390
      :
      :
      :
QY 479 GGCAGAAAGACGATCCGCTCATCGGGTGGTGTCTCTCTCATTTGCCCTGGTGGTTTC 538
      :
      :
      :
Db 391 -----ThrSerLeuAlaAlaSerSerSerSerHisIleTyrProSerSe 404
      :
      :
      :
QY 539 CTATCATCTCTTCCAGTTCTGCGAGGGCCACACAGGATACAGTACAGTACAGAGAG 598
      :
      :
      :
Db 404 rThrThrSerThrThrSerSerThrThrSerThrThrThrThrThrThrThrArgAr 424
      :
      :
      :
QY 599 GAGAGCTGTCCCAAGACAGCTGTGCTGCTGTGACCGGGGTGTGACATGCAAGTGAAGT 658
      :
      :
      :
Db 424 gThrThrThrThrProThrThr-----ThrThrArgArgHisThrThrThrThrThr 440
      :
      :
      :
QY 659 GACGAGCTGGAGCTGAGAGTTGACTGGGACAAAGTCTGCTTAAATCTTACTCTGGG 718
```

```

Db      440 oThArgProtyrGlnArgProthrAlaThrSerSerSerthrThr----- 458
Qy      719 TCCGCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGCAATGACTCTACTCAGAGAG 778
Db      459 -----ThSerSerSerthrThrThrThrThr 468
Qy      779 ACCTGACGAGCTGGGTTTCGAG-----AGTGTCACCGGAGAAC 819
Db      468 GPrIILSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 488
Qy      820 CGAGGTGCCCAAGGATTTTGCCAAAGCTTCTCAATCTTGAGATTAACAATCCACCAT 879
Db      488 rGlnProThrHisArgThrProValLeuAlaThrSerGlyLeuGlnThrAnGlnIleSe 508
Qy      880 CCAGAAAGCTCCACAGCTGTAATGCCCTCCAG---CGTAATCTCTCCACAGT 936
Db      508 rAspSerSerLeuProAspAlaGlyAlaLeuGlyArgValThrIleSerAlaAlaAr 528
Qy      937 TTCCCACTGGGAGTGAAGGCCCATGACC-----GGGGGAGTCGTGGAGGGGGCT 987
Db      528 gSerGlyCysGlyAlaProThrLeuAlaAlaArgProGlnThrArgIleValGlyLySe 548
Qy      988 GGCTTCGATATGACAGTGGCTTGGCAAGTGAATCTGCAC-----TT 1029
Db      548 rAlaAlaPheGlyArgTrpProTrpGlnValSerValArgArgThrSerPhePheGlyPh 568
Qy      1030 CGGACCAACCCACATCTGTGGAGGCGCTGATTCAGCGCCAGTGGGTCTACTGCCGC 1089
Db      568 eSerSerThrHisArgCysGlyGlyAlaLeuIleAsnGlnAsnTrpIleAlaThrAlaG 588
Qy      1090 CCACTGC-----TTCTTCGTAACCCCGGAGAAAGTCTCTGAGAGGCTGGAAGTGA 1140
Db      588 yHIScysValAspAspLeuLeuIleSeGlnIleArgGlyLeArgValGly-----GluTy 606
Qy      1141 CGCGGGACGACGACCACTGACACGATTGCCGAG---GCAAGCTCCATTCGCGAGATCAT 1197
Db      606 rAspPheSerHisValGlnGlnGlnLeuProTyrlleGlnArgIleValAlaAlaValSva 626
Qy      1198 CATCAACAGCAATTAACCGATGAGAGAGACGACTATGACATCGCCCTCATGCGGCTGT 1257
Db      626 lValHisProIleArgSerPheLeuThrTyrlGluTyAspLeuAlaLeuValLySeGln 646
Qy      1258 CAAGCCCTTGAACCTGTCGCTCATCCACCTCTGCTTCCCTCCCATGATGACAGAC 1317
Db      646 uGlnProLeuGlnPheAlaProHisValSerProIleCysLeuPro-----GluTh 663
Qy      1318 CTTAGGCTC-----AATGAGACCTGTGATCAACAGCTTTGGCAGACGAG 1365
Db      663 rAspSerLeuLeuIleGlyMetAlaAlaThr-----ValThrGlyTrpGlyArgLeuSe 681
Qy      1366 GGAGACAGATGACAAGACATCCCTTCCTCCGAGAGGTGACGATCATCATGACTT 1425
Db      681 rGlnGlyGlyThrLeuProSer---ValLeuGlnGlnValSerValProIleValSerAs 700
Qy      1426 CAGAATAATGCATATCACTTGTGTC-----TATGACAGTTAATCTTACCCCAAGAT 1476
Db      700 nAspAsnCysIleSerMetPheMetArgAlaGlyArgGlnGlnPheIleProAspIlePh 720
Qy      1477 GATGTGTCGGGAGACCTGCTGGGGGAGAGATCCCTGCGAGGAGAGACGGGGGGCC 1536
Db      720 eLeuGlyAlaGlyIleGlyThrGlyGlyGlnAspSerCysGlnIleAspSerGlyGlyPr 740
Qy      1537 TCTTGTCTGTAG---CAGAACCAACCGGTGTACTGTGACAGTGTCAACGCTGGGAC 1593
Db      740 oLeuGlnAlaIleAspSerGlnAspGlyArgPhePheLeuAlaIleIleSerTrpGlyI 760
Qy      1594 AGGTGTGGCCAGGAACAACCTGTGTGTATACCAAAATGACAGAAGTCTTCCCTG 1653
Db      760 eGlyCysAlaGlyAlaAsnLeuProGlyValCysThrArgIleSerLySphenTrpTr 780
Qy      1654 GATT 1657

```

```

Db      780 pile 781
RESULT 13
KFBI1
N:Alternate names: antithrombophilic factor C; plasma thromboplastin antecedent
C/Species: Homo sapiens (man)
C/Date: 13-Aug-1986 #sequence revision 26-May-1994 #text_change 05-Oct-2004
R/Accession: A27431; A00920; A37940
R/Akaaki, R.; Davie, E.W.; Chung, D.W.
Biochemistry 26, 7221-7228, 1987
A>Title: Organization of the gene for human factor XI.
A:Reference number: A27431; MUID:88107663; PMID:2827746
A:Accession: A27431
A:Molecule type: DNA
A:Residues: 1-625 <AS>
A:Cross-references: UNIPROT:P03951; UNIPARC:UPI000000DB87; GB:M18295
A/Note: the sequence shown follows the authors' translation
R/Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
Biochemistry 25, 2417-2424, 1986
A>Title: Amino acid sequence of human factor XI, a blood coagulation factor with four tar
A:Reference number: A00920; MUID:86243360; PMID:3636155
A:Accession: A00920
A:Molecule type: mRNA
A:Residues: 1-625 <FU>
A:Cross-references: UNIPARC:UPI000000DB87; GB:M13142; NID:G182832; PIDN:AA52487.1; PID:G
R/Mueller, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2056-2060, 1991
A>Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
A:Reference number: A37940; MUID:91152017; PMID:1998667
A:Accession: A37940
A:Molecule type: protein
A:Residues: 28-33;35-49, 'X', 51-55, 'X', 57-63;70-75, 'X', 77-79;107-109, 'X', 111-112;132-139, '
;280-282, 'X', 284;285-297;313-316, 'X', 318-319;320-326, 'X', 328-330; 'X', 347-349;373, 'X', 375;
A:Cross-references: UNIPARC:UPI00000172B4; UNIPARC:UPI00000172B5; UNIPARC:UPI00000172B6;
B2B; UNIPARC:UPI00000172B2C; UNIPARC:UPI00000172B2D; UNIPARC:UPI00000172B2E; UNIPARC:UPI000
10000172B3A; UNIPARC:UPI00000172B3; UNIPARC:UPI00000172B36; UNIPARC:UPI00000172B37; UNIPARC
C/Comment: The proenzyme consists of two identical chains linked by one or more disulfide
he active site, and a heavy chain, which associates with high molecular weight (HMW) kind
C/Genetics:
A/Gene: GDB:F11
A/Cross-references: GDB:119891; OMIM:264900
A/Map position: 4q35-4q35
A/Intons: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 526
C/Function:
A:Description: catalyzes the proteolytic activation of coagulation factor IX
A/Pathway: blood coagulation intrinsic pathway
C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C/Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-387/Product: coagulation factor XIa heavy chain #status experimental <HCH>
F.19-108/Domain: apple repeat <AP1>
F.109-198/Domain: apple repeat <AP2>
F.199-288/Domain: apple repeat <AP3>
F.290-379/Domain: apple repeat <AP4>
F.388-625/Product: coagulation factor XIa light chain #status experimental <LCH>
F.388-618/Domain: trypsin homology <TRY>
F.20-103;514-581;571-599/Disulfide bonds: #status predicted
F.29/Disulfide bonds: interchain #status experimental
F.146-76;50-56;110-19;136-165;140-146;200-283;226-255;230-236;291-374;317-346;321-327;386
F.339/Disulfide bonds: interchain #status predicted
F.388-388/Cleavage site: Arg-Ile (coagulation factor XIa) #status experimental
F.431;480;575/Active site: His, Asp, Ser #status predicted
F.491/Binding site: carbohydrate (Asn) (covalent) #status experimental
Alignment Scores:
Pred. No.: 4,54e-22 length: 625
Score: 532.00 Matches: 138
Percent Similarity: 45.5% Conservative: 62
Best Local Similarity: 31.4% Mismatches: 136
Query Match: 16.0% Indels: 104
DB: 1 Gaps: 17

```

```

US-10-806-370-11 (1-1748) x KFHUI (1-625)
Oy 609 CCAAGCAGCGCTGTTGAGCGGCGGCTGAGCTGCAAGCTGAAGAGTGAAGCTG- 667
Db 206 ProanthraValPheAla--AAserAenILeAAserValMeAlaProAAspAlaPheV 225
Oy 668 -----GGCTGCGTGAAGTTT-----G 683
Db 225 aLcyGgLYArgILeCyethrHISpRogLYcyleuPhePheThrPhePheSerGing 245
Oy 684 ACTGGAGCAAGTCT-----CTGCTTAATAATCTACTCTGGCTCTCC 725
Db 245 IutPrProLYgLYserGlnArgAenLeuCYLeuLeuLYs-----ThrsArg 261
Oy 726 ATGAGTGCTTCCC-----ATCTGAGCAAGCACTGAATGACTCTTACAGAGA 776
Db 261 IuserGlyLeuProSerThrArgILeLYserLYsValLeuSerGlyPheSerLeug 281
Oy 777 AGACCTGCAGCAGCTGGGTTTCAGAGTGTCTACCGGCAACCGAGGTTCCGACAGG 836
Db 281 InserCYArg-----HiserILeProValPheCYSHSerS 294
Oy 837 ATTTGCCACAGCTTCTCAATCTTGAGATCAACTCCACATCCAGAGAACTCCACA 896
Db 294 erPheTYrHISpRThrAAspRhePheLeuGlyGlnIuLeuAAspILeValAlaLYsSerH 314
Oy 897 GGTCTGAATGC-----C 908
Db 314 IeGluAlaCYgGLnLYleuCYeThrAenAlaValArgCYgGLnPhePheThrTYrThP 334
Oy 909 CTTCACGCGG-----TATATCTCCCTCGAGTGTCC 941
Db 334 roAlaGlnAlaSerCYsAenGlnLYgLYsCYeTYLeuLYleuLYserSerAeng 354
Oy 942 ACTGCGGACTGAGGCGCATGACCGG----- 967
Db 354 IYserProThrLYleuHISgLYArgGlyLYleSerGlyTYrThrLeuArgLeuC 374
Oy 968 -----CGGATCGTGGAGGCGGCTGG 989
Db 374 yelYMetAAspAenGlyCYeThrThrLYleLYsPProArgILeValGlyLYThraAs 394
Oy 990 CCTCGATAGCAGTGGCTTGGCAAGTGAAGTGTGCTGCGGACACAC----- 1039
Db 394 erValArgGlyGlnTrpProTrgInValThrLeuHIS-----ThrThrserProThrg 412
Oy 1040 -----CACATCTGTGAGAGCAGCTCATTAAGCCGAGTGGTGTCTACCTGCCGCCACT 1094
Db 412 InArgHISLeuCYgLYserILeILeGlyAenGlnTrpILeLeuThraAlaHIS 432
Oy 1095 GCTTCTTC---GTGACCGCGGAGAAAGCTCGAGGCGTGAAGTGAAGCGGAGCA 1151
Db 432 yPheTYrGlyValGlnIuSerProLYsILeLeu-----ArgValTYserGlyLYleL 449
Oy 1152 GCAACCTGCACAGCTTGCCTGAGGAGGCTCC-----ATTGCCAGATCATCATCA 1202
Db 449 euAenGlnInserGlyILeLYgLYuAAspRhePheGlyValGlnIuILeILeH 469
Oy 1203 ACAGCAATTACCGATGAGAGACACACTGATGACCTGAGCTGAGCTGCGCTTCCAAAGC 1262
Db 469 IeAAspGlnTYrLYMetAlaGlnIuSerGlyTYrAAspILeAlaLeuLYleuGlnIuThT 489
Oy 1263 CCTGACCGCTGCGCTGATCCACCTGCTGCTCCGATCGGATCGAGACACTTTA 1322
Db 489 hrValaenTYrThraAspserGlnArgProILeCYsLeuProSerLYsGlyAAspArgAenV 509
Oy 1323 GCCTCAATGAGACCTGCTGATCAACAGGCTTGGCAAGACGAGGAGACAGATGACAAGA 1382
Db 509 alIeTYrThraAspCYeTrpValThrGlyTYrGly---TYArgLYsLeuArgAAspLYsI 528
Oy 1383 CATCCCTTCTCGCGGAGGTGACAGGTCAATTCATGACTGACCTCAAGAAATGCAATGACT 1442

```

```

Db 528 IeGlnAenThrLeuGlnLYleAlaLYsILeProLeuValThraenGlnIuCYeGlnLYsA 548
Oy 1443 ACTTGCTATGACAGTTAC---CTTACCCCAAGATGATGTGCTGGACCTTGCTG 1499
Db 548 rg-----TYArgGlyHISLeLYleThrHISLYsMetILeCYAlaGlyTYrArgGln 566
Oy 1500 GGGGCAAGACTCTGCGCAGGAGACAGCGGGGCGCTTGTCTGTGAGCAAGCAACC 1559
Db 566 IYgLYsAAspAlaCYLeGlyAAspserGlyLYPProLeuSerCYeLYHISaenGlnV 586
Oy 1560 GCTGTACTCTGGCAGGTGTCAACGAGTGGGCAAGAGCTGTGGCCAGAGAAACAACCTG 1619
Db 586 alTrpHISLeuValGlyILeThrserTrpGlyGlnLYsAlaGlnArgIuArgProG 606
Oy 1620 GTGTGTACCAAGTGTCTCTCCCTGATTTACAGCAAGATGAGAGAC 1675
Db 606 IYValTYrThraenValaGlnIuTYrValaAspTrpILeLeuGlnIuLYserThraAla 624

RESULT 14
PLBO
plasma (BC 3.4.21.7) precursor - bovine
N.Alternate names: plasminogen
C.Species: Bos primigenius taurus (cattle)
C.Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C.Accession: S45046; A25835; I45961; S03736
R.Berglund, L.; Andersen, M.D.; Petersen, T.B.
Submitted to the EMBL Data Library, May 1994
A.Description: Cloning and characterization of the bovine plasminogen cDNA.
A.Reference number: S45046
A.Accession: S45046
A.Molecule type: mRNA
A.Residues: 1-812 <BER>
A.Cross-references: UNIPROT:P06868; UNIPARC:UPI0000043BEV; EMBL:X79402; NID:G494962; PID:
A.Experimental source: liver
A.Note: it is uncertain whether Met-1 or Met-8 is the initiator
R.Schaller, U.; Moser, P.W.; Damminger-Mueller, G.A.K.; Rosselet, S.J.; Kamper, U.; Rickl
Eur. J. Biochem. 149, 267-278, 1985
A.Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmi
A.Reference number: A25835; MUID:85203906; PMID:3846532
A.Accession: A25835
A.Molecule type: protein
A.Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
A.Cross-references: UNIPARC:UPI0000172B95
R.Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A.Title: Characterization of a complementary deoxyribonucleic acid coding for human and t
A.Reference number: I45961; MUID:85023311; PMID:6148961
A.Accession: I45961
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 706-743, 'R', 745-812 <MAL>
A.Cross-references: UNIPARC:UPI000016C365; GB:K02935; NID:G163551; PID:AAA30714.1; PID:
R.Brundholz, R.A.; Lerch, P.G.; Schaller, U.; Rickl, E.E.; Leisler, W.; Mannborg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A>Title: Comparison of the primary structure of the N-terminal CNBr fragments of human, t
A.Reference number: S03735; MUID:81212097; PMID:7238497
A.Accession: S03736
A.Molecule type: protein
A.Residues: 27-83 <BRU>
A.Cross-references: UNIPARC:UPI0000172B96
C.Function:
A.Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
in the walls of the graafian follicle; also activates the urokinase-type plasminogen acti
A.Pathway: fibrinolysis
C.Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology;
C.Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma;
F:1-66/Domain: signal sequence #status predicted <Sig>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:27-812/Product: plasminogen #status experimental <PRO>
F:27-103/Domain: activation peptide #status experimental <APT>
F:104-583, 584-812/Product: plasmin chain A #status experimental <AAT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringle homology <KRI>

```


F_192-269/Domain: kringle homology <KR2>
F_282-359/Domain: kringle homology <KR3>
F_384-461/Domain: kringle homology <KR4>
F_485-564/Domain: kringle homology <KR5>
F_584-812/Domain: plasmin chain B #status experimental <BCH>
F_584-805/Domain: trypsin homology <TRY>
F_56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,333-364,365-400/Domain: #status predicted
F_315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F_365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F_624,667,762/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	1,6e-21	length:	812
Score:	522.00	Matches:	168
Percent Similarity:	40.7%	Conservative:	67
Best Local Similarity:	29.1%	Mismatches:	216
Query Match:	15.7%	Indels:	127
DB:	1	Gaps:	24

US-10-806-370-11 (1-1748) x PLBO (1-812)

OY	10	CATGACGAGGAGACAGCCAGCGGAATGCAATCTCCACAGAAACACCTTACGTGG-----	63
Db	335	AsnGlyValuysAlaPro---TrpCysTrpThrTrnAsnSerGluValAlaGTrpGluTr	353
OY	64	---AGCATCTCCAGCCACGAGCATCTCCAGCTGGGACACCTCCAGCGCGGAGCATCTCCAGC	120
Db	354	CysThrIleProSerGysGluSerSerProIeuSerThrGluAlaGmeAlaProVal	373
OY	121	CCAGGCACTCCAGCCCA-----GGCATCTCCAGCTGG	153
Db	374	ProProGluGlnThrProValProGlnAlaProCysTrpTrnIsgIAsnGlyGlnSerTrpArg	393
OY	154	GAACCTCCGGGGCGGGGCGGCTCCAGCCAGCCAGCATCTCCAGCTGTATACCTCCAGGCGG	213
Db	394	GlyThrSerSerThrThrIleThrGluArgIysCysGlnSerTrpSerSerMetThrPro	413
OY	214	GGCATCTCCAGCGCGGAGCATCTCCAGCCAGGCACTCCAGCCCGGAGCATCTCCGGCTCT	273
Db	414	His-----	414
OY	274	GGCATCACTTCCAGGCTCTCATCCGGAGTCATCATCCGCCAGTGACGCTCGTGAC	333
Db	415	-----ArgHisIeuLysThrProGlu	421
OY	334	AACCTCCCAACCAAGATGTAACCTGTTTAGAACAAACCAAGTGGGGCTGTACCAATCCG	393
Db	422	AsnTrpProAsnAlaIeuLys-----ThreAsnTrpCysArgAsnPro	436
OY	394	ATCATCTCCGCGCAGGTGACAGCCAGCAACGAGGCCACAGGAGAGCCGAGGTACGAG	453
Db	437	-----AspAlaAspLysSerProTrpCysTrpThrTrnAspProAlaValArg	452
OY	454	CTGCCCAGATTCAACCTGCGGGAGGGGCGGACAGAGCATCCGCTCATCGGGTCCGTCT	513
Db	453	TrpGluPheCysAsnIeuLysIysCysSerGlu-----ThrProGluGlnValProAla	470
OY	514	CTCTCTCATTTGCCCTGTGGTTTCGCTATCATCTCTTCCAGTTTGGCAGGGCCACAC	573
Db	471	AlaProGlnAlaIleProIly-----ValGluAsnPro-----ProGlu	482
OY	574	AGGATCAGGTATCAAGAGAGGAGGAGAGATGTGCCAAGCAGCGTGTGCTGTGACGG	633
Db	483	AlaAspCysMetIleGlyThrIleGlyLysSerTrpArgGlyLysIysAlaThr-----Thr	500
OY	634	GGTGTGGATCTGCAGACTGAAGAATGACAGACTGGGCTGCGTGAAGTTTGACTGGACAA	693
Db	501	ValAlaGlyValProCysGlnGlu-----	508
OY	694	GTCCTGCTTAAATATCTACTCTGGGGCTCCCATCATAGTGGCTTCCCATCTGTAGACAA	753
Db	509	-----TrpAlaIaGlnGluProHisGlnIleSer-----	518

QY	754	CTGAATGACTCCTACTACAGAGAAAGACTCCAGACACTGGCTTTCAGAG-----TGC	807
Db	519	-----IIlePheThrProGluThrAsnProGlnSerGlyLeuGluArgAsnIleTyrCys	535
QY	808	---TACCCGGCAACCCAGGTTGGCCCAAGGGA---TTTGGCCAAACAGCTTCTGAATCTT	861
Db	536	ArgAsnProAsp-----GlyAspValAsnGlyProTyrCysTyrThrMet-----	550
QY	862	GAGATACAACTCCACCATCCAGAAAGCTTCACAGGCTTGAAATGCCCTTTCAGCGGTA	921
Db	551	-----AsnProArgLys-----ProPheAspTyrCys	559
QY	922	TATTCCTCCTCAGTGTTC-----CACTGGGACTG-----AGGC	957
Db	560	AspValPro-GlnCysGluSerSerPheAspCysGlyLysProLysValGluProLys	579
QY	958	CATGACCGGGGCGATCGTGGGAGGGGCGCTGGCTCGGATGACAAAGTGGCTTGGCAAGT	1017
Db	579	cySerSerGlyArgLleValGlyGlyCysValSerLysProHiserThrProTrrGlnVa	599
QY	1018	GAGCTTCGACTTCGGCACCCACCATCTGTGAGGACCGCTCAATTGACGCCAGTGGT	1077
Db	599	IserLeuArgArgSerSerArgHisPheCysGlyGlyThrLeuIleSerProLysTrrVa	619
QY	1078	GCTTACTGCGCCGCACTGCTTCTTCGTGACCCCGGAAAGGCTCTGGAGGGCTGGAAGT	1137
Db	619	IleuThrAlaIleHisCys-----LeuAspAsnIleLeuAlaLeuSerPheTyrLysVa	637
QY	1138	GTAGCGCGGCGCACGAACTGTCACAGTTGGCTTGAGGACGCTCCCAATGCGCAGATCAT	1197
Db	637	IleLeuGlyValHisAsn-----GluLysValArgGluGlnSerValGlnGluIlePr	655
QY	1198	CATCAACAGCAATTTACACCGATGAGGAGAGAGCATATGACATCGCCCTGATCGGCTGTC	1257
Db	655	oval---SerArgLeuPheArgGluProSerGlnAlaAspIleAlaLeuLeuLysLeuSe	674
QY	1258	CAGCGCCCTGACCTGTGCGCTCACAATCCACCTGCTTGGCTCCCTCCCATGATGAGACAGAC	1317
Db	674	TArgProAlaIleIleThrLysGluValIleProAlaCysLeuProProProAsnTyrMe	694
QY	1318	CTTAAAGCTCAATGAGACCTGTGCTGATCAACAGCTTGTGGCAAGACCAAGGACGATGA	1377
Db	694	tValAlaIleArgThrGluCysTyrIleThrGlyTrrGlyGluThrGlnGlyThrPheG1	714
QY	1378	CAAGACATCCCTCTCTCCGGGAGGCGAGTCAATCTTCATCTGACATCAAGAAATGCAA	1437
Db	714	ygLu-----GlyLeuLeuLysGluAlaHisLeuProValIleGluLysnLysValCysAs	732
QY	1438	TGACTTACTGGTCTATGACAGTTACTTTACCCCAAGATGATGTGTGTGGGACCTTGC	1497
Db	732	naArgAsnGluTyrLeuAspGluArgValLysProThrGluLeuCysValGlnHisLeuI1	752
QY	1498	TGGGGGAGAGACTCCCTGGCAGGAGGAGGAGGGGGGCTTGTCTGTGACGAGAACAA	1557
Db	752	eGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAs	772
Db	772	plysTyrIleLeuGlnGlyValThrSerTrrGlyLeuGlyCysValArgProAsnLysPr	792
QY	1618	TGGTGTGTACCAACAAAGTACAGAAATCTTTCCTCGATTTTACAGCAAGATG	1669
Db	792	oGlyValTyrValArgValSerProTyrValProTrrIleGluGluThrMet	809

RESULT 15

```

A57014
prolactin (EC 3.4.21.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2000
C:Accession: A57014; A54866
R:Yu, J.X.; Chao, L.; Chao, J
S: Biol. Chem. 270, 13483-13489, 1995

```


Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rnp**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 01:24:21 | Search time 106.6 Seconds
(without alignments)
4550.452 Million cell updates/sec

Title: US-10-806-370-11

Perfect score: 3333
Sequence: 1 ctccagagaccatggagaggg.....ggcgcgcgtgactcggagaaa 1748

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 5699196

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWB/spool/US10806370/runat.15092006.105625.9965/app.query.fasta-1
-DB=UniProt -QEXT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=humana0.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORER=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ds802p
-USER=US10806370 @CGN 1 1 891 @runat.15092006.105625.9965 -NCPU=6 -ICPU=3
-NO MMAP -NCG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: UniProt_7.2.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2956	88.7	581	1	TMPSD_HUMAN
2	2510.5	75.3	543	1	TMPSD_MOUSE
3	891	26.7	359	2	Q4SPG0_TETNG
4	784.5	23.5	484	2	Q311V4_PANPA
5	783.5	23.5	484	2	Q311V3_PPRIM
6	780.5	23.4	490	1	Q311V5_PANTR
7	777.5	23.3	490	1	TMPS2_MOUSE
8	774.5	23.2	484	2	Q311U4_MACMU
9	773	23.2	490	2	Q6P7D7_RAT
10	771.5	23.1	490	2	Q3UKB3_MOUSE
11	770.5	23.1	490	2	Q7NN04_MOUSE
12	768.5	23.1	492	2	Q6GKT7_HUMAN
13	767.5	23.0	484	2	Q311U8_HYUSY
14	766.5	23.0	492	1	TMPS2_HUMAN
15	766.5	23.0	492	2	Q96T73_HUMAN
16	764	22.9	490	2	Q920K3_RAT

17	763.5	22.9	453	2	Q812A6_MOUSE	Q812A6 mus musculus
18	763.5	22.9	453	2	Q3T206_MOUSE	Q3T206 mus musculus
19	762.5	22.9	475	1	TMPS3_MOUSE	Q8K10 mus musculus
20	762.5	22.9	453	2	Q2M1G4_MOUSE	Q2M1G4 mus musculus
21	759.5	22.8	486	2	Q5PR46_BRARE	Q5PR46 brachydanio
22	736.5	22.1	538	2	Q5USC7_HUMAN	Q5USC7 homo sapien
23	731	21.9	454	1	TMPS3_HUMAN	P57727 homo sapien
24	713.5	21.4	722	2	Q6NUP5_XENLA	Q6NUP5 xenopus lae
25	713	21.4	767	2	Q9DCR2_XENLA	Q9DCR2 xenopus lae
26	697.5	20.9	437	1	TMPS4_HUMAN	Q9N154 homo sapien
27	696	20.9	445	2	Q8C1U7_RAT	Q8C1U7 ratius norv
28	689.5	20.7	439	2	Q5RDX7_PONPY	Q5RDX7 pongo pygma
29	684.5	20.5	435	1	TMPS4_MOUSE	Q8VC45 mus musculus
30	684	20.5	388	2	Q4RRR7_TETNG	Q4RR7 tetradon n
31	668.5	20.1	371	2	Q8C1U6_RAT	Q8C1U6 ratius norv
32	664	19.9	455	1	TMPS5_MOUSE	Q9E104 mus musculus
33	659	19.8	455	2	Q8CDR0_MOUSE	Q8CDR0 mus musculus
34	644.5	19.3	457	1	TMPS5_HUMAN	Q9H363 homo sapien
35	620.5	18.6	445	2	Q3U0U6_MOUSE	Q3U0U6 mus musculus
36	612.5	18.4	436	1	HEPS_MOUSE	Q35453 mus musculus
37	611.5	18.3	417	1	HEPS_HUMAN	P05981 homo sapien
38	609.5	18.3	417	2	Q5R5E8_PONPY	Q5R5E8 pongo pygma
39	600	18.0	326	2	Q7Z280_BRARE	Q7Z280 brachydanio
40	592.5	17.8	416	1	HEPS_RAT	Q05511 ratius norv
41	589.5	17.7	572	1	TMPS7_MOUSE	Q8B1K6 mus musculus
42	585.5	17.6	829	2	Q4PCJ3_MOUSE	Q4PCJ3 mus musculus
43	582.5	17.5	730	2	Q4RH0_TETNG	Q4RH0 tetradon n
44	580	17.4	1019	1	ENTK_HUMAN	P98073 homo sapien
45	580	17.4	1019	2	Q2NKL7_HUMAN	Q2NKL7 homo sapien

ALIGNMENTS

RESULT 1
TMPSD_HUMAN STANDARD; PRT; 581 AA.
ID Q9BYE2; Q86YH4; Q96JY8; Q9BYE1;
AC 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 15-MAR-2005, sequence version 2.
DT 07-MAR-2006, entry version 29.
DE Transmembrane protease, serine 13 (BC 3.4.21.-) (Mosaic serine
protease) (Membrane-type mosaic serine protease).
OS Homo sapiens (Human)
GN Name=TMPS13; Synonym=MSP, TMPS11;
NC Name=sapien (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [mRNA] (ISOFORMS 1 AND 3), AND TISSUE SPECIFICITY.
RC TISSUE=Lung; PubMed=11267681; DOI=10.1016/S0167-4781(01)00184-1;
RX Kim D.R., Sharmin S., Inoue M., Kido H.,
RA "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [mRNA] (ISOFORM 2).
RA Park T.J., Park W.J.,
RT "Homo sapiens transmembrane protease, serine 6 (TMPS6) mRNA."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA] (ISOFORM 4).
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Tanaka H., Yasuda T., Watanabe T., Watanabe M., Shiraori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,

RA Ntmonlyu K., Ichihashi T., Yameashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimeta M., Watanabe M., Hiraoka S., Chiba Y., Ichida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houcha T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togeishi T., Komai F., Hara R., Takuchi K., Arita M., Imose N.,
RA Masehino K., Yuuki H., Oshima A., Sasaki N., Aocuka S.,
RA Yoshikawa Y., Matsumura H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Tereshima Y., Suzuki O.,
RA Nakagawa S., Senon A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumezaki A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohnori Y.,
RA Kawabata A., Hikiji T., Kobutake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togeishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mutsushima-Sugano J., Satoh T., Shirai T., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kituchi H., Maeno Y., Yamashita R.,
RA Nakai K., Yida T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs." ;
RA Nat. Genet. 36:40-45(2004) .
CC [4]
RN NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RP NCBI
RG Human chromosome 11 International sequencing consortium;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=MSBP, Large form, Membrane-type;
CC IsoId=Q9BYE2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9BYE2-2; Sequence=VSP_013103; VSP_013104;
CC Note=No experimental confirmation available;
CC Name=3; Synonyms=MSBP, Small form;
CC IsoId=Q9BYE2-3; Sequence=VSP_013099, VSP_013102;
CC Name=4;
CC IsoId=Q9BYE2-4; Sequence=VSP_013100, VSP_013101;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Isoform 1 and isoform 3 are predominantly
expressed in lung, placenta, pancreas, and prostate. Isoform 3 is
weakly expressed in testis and peripheral blood lymphocytes.
CC -1- POLYMORPHISM: The repeat A-8-P-A-[GQR] is polymorphic and the
number of copies varies between 12 to 14.
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDI-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -1- SIMILARITY: Contains 1 SRCR domain.
CC -1- CAUTION: Ref.2 has referred to this protein as TMPRSS6.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License

DR	PRINTS; PRO0723; CHYMOTRYPSIN.
DR	SMART; SMO0020; TRY_Spc; 1. FALSE_NEG.
DR	PROSITE; PS01209; LIDRA_2; FALSE_NEG.
DR	PROSITE; PS00068; LIDRA_1; FALSE_NEG.
DR	PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR	PROSITE; PS0287; SRCR_2; 1.
DR	PROSITE; PS0240; TRYPsin DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; 1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Alternative splicing; Glycoprotein; Hydrolyase; Membrane; Polymorphism;
KW	Protease; Repeat; Serine protease; Signal-anchor; Transmembrane.
FT	CHAIN 1 581 /FTid=PRO_000008698
FT	TOPO_DOM 1 160 Cytoplasmic (Potential).
FT	TRANSMEM 161 181 Signal-anchor for type II membrane protein (Potential).
FT	EXTRACELL 182 581 Extracellular (Potential).
FT	REPEAT 9 13 1-1.
FT	REPEAT 14 18 2-1 (approximate).
FT	REPEAT 19 23 1-2.
FT	REPEAT 24 28 1-3.
FT	REPEAT 29 33 2-2.
FT	REPEAT 34 38 1-4.
FT	REPEAT 39 43 1-5.
FT	REPEAT 44 48 1-6.
FT	REPEAT 49 53 2-3.
FT	REPEAT 54 58 1-7.
FT	REPEAT 59 63 1-8.
FT	REPEAT 64 68 2-4.
FT	REPEAT 69 73 1-9 (approximate).
FT	REPEAT 74 78 1-10.
FT	REPEAT 79 83 1-11.
FT	REPEAT 84 88 1-12.
FT	SRCR 190 320 SRCR.
FT	DOMAIN 199 221 LDL-receptor class A.
FT	DOMAIN 321 554 Peptidase S1.
FT	REGION 9 88 12 x 5 AA repeats of A-S-P-A-[GLQR].
FT	REGION 14 89 4 x 5 AA repeats of T-P-G-R.
FT	COMPBIAS 9 68 Ala-rich.
FT	ACT_SITE 361 361 Charge relay system (By similarity).
FT	ACT_SITE 409 409 Charge relay system (By similarity).
FT	ACT_SITE 506 506 Charge relay system (By similarity).
FT	CARBOHYD 250 250 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD 287 287 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
FT	CARBOHYD 440 440 N-linked (GlcNAc...) (Potential).
FT	DISULFID 245 309 By similarity.
FT	DISULFID 258 312 By similarity.
FT	DISULFID 346 362 By similarity.
FT	DISULFID 443 512 By similarity.
FT	DISULFID 475 491 By similarity.
FT	DISULFID 502 530 By similarity.
FT	VARSPLIC 146 181 GTSLPKPTMRGGQQLPLIGVLLILALVSLILF -> V (in isoform 3).
FT	VARSPLIC 423 486 /FTid=VSP_013099.
FT	VARSPLIC 486 581 AHIRPACIPMKGFIFSLNETCWITFGKTREREDTKTSPFLAREVOVALIDFKKNLYVDLSLT -> GEIGCTPRSPAPOPCHPLDPHLSSAVNSYVPGRKASAGCKSKTLKDPMHEHFCPIIRETEAQL (in isoform 4).
FT	VARSPLIC 581 581 /FTid=VSP_013101.
FT	VARSPLIC 581 581 VRSLDDPAISRLGTSSGCGPGCAPRV -> SEVERFKS (in isoform 3).
FT	VARSPLIC 581 581 /FTid=VSP_013102.
FT	VARSPLIC 581 581 VRSL -> SSAG (in isoform 2).
FT	VARSPLIC 581 581 /FTid=VSP_013103.
FT	VARSPLIC 581 581 Missing (in isoform 2).
FT	VARSPLIC 581 581 /FTid=VSP_013104.
FT	VARSPLIC 581 581 H -> Y (in Ref. 2).
FT	VARSPLIC 581 581 Y -> H (in Ref. 1); BAB39741).
FT	VARSPLIC 581 581 K -> E (in Ref. 2).
FT	VARSPLIC 581 581 C -> R (in Ref. 2).

FT CONFLICT 259 259 Q -> R (in Ref. 1).
 FT CONFLICT 298 298 E -> H (in Ref. 1).
 FT CONFLICT 496 496 H -> R (in Ref. 2).
 SQ SEQUENCE 581 AA; 62680 MM; A4935CCCP31D29EE CRC64;

Alignment Scores:

Prod. No.: 3.66e-158 Length: 581
 Score: 2956.00 Matches: 553
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 88.7% Indels: 0
 DB: 1 Gaps: 0

US-10-806-370-11 (1-1748) x TMPSD_HUMAN (1-581)

QY 11 ATGAGAGGAGCAACGCGGAGATGATCTTCAGCAAGAACACCTTACGTGAGCATCT 70
 Db 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAAGCCAGGCATCTCAGCTGGGAGCACTCCAGCGCGGAGATCTCCAGCCAGGCATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAAGCCAGGCATCTCAGCTGGGAGCACTCCAGCGCGGAGATCTCCAGCCAGGCATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCTGGTACACCTCCAGGCGCGGAGCATCTCCAGCGCGGAGATCTCCAGCCAGGCATCT 250
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAAGCCCGGCATCTCCGGCTTCGGCATCTTCCAGGTCCTCATCCGGCAGGTCATCA 310
 Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerLeuSerSerSerSerGlyArgSerSer 100
 QY 311 TCCGCGAGCTCAGCGCTGGTGAACAACCTCCCAACAGAGTGAACCTTGTGTAAGACAAC 370
 Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValAlaGlnAlaThr 120
 QY 371 CCAAGTGGGGCTGTACCCATCCGATCATCTTCGCGAGTGCAGCAACAGCAACAGGACC 430
 Db 121 ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGCCAGGTAAGAGCTCCGCAAGTTCACTCGCGGAGGCGGAGGCGAACAAG 490
 Db 141 ThrArgIleSerProGlyThrSerLeuProIleThrArgGlyGlnGlyGln 160
 QY 491 CTACCGCTCATCGGGTGGTGGCTCTCTCATCTGCGCTGGTGGTGGTGGTGGTGGTGGT 550
 Db 161 LeuProIleIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleLeu 180
 QY 551 TTCAGATTCTGGCAGGAGCAACAGGAGTACAGTACAGAGCAGAGAGAGCTGTGCC 610
 Db 181 PheGlnIleThrArgGlnGlnIleThrGlyIleArgGlyTyrGlnGlnIleArgGlnCysPro 200
 QY 611 AAGACGCTGTCTGCTGTACGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 670
 Db 201 LysHisAlaValAlaGlyCysAspGlyValValAspCysLysLeuLysSerAspGlnLeuGly 220
 QY 671 TGGGTGAGGTTTGAAGTCAAGTCTCTGTTAAATCTACTCTGGGCTCTCCATCAG 730
 Db 221 CysValAlaGlnPheAspIlePheLysSerLeuLeuLysIleTyrSerGlySerSerHisGln 240
 QY 731 TGGGTTCCATCTGTGACAGCAACTGGAATGACTCTTACTCAGAGAAAGCTGCGACAG 790
 Db 241 TrpLeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlyLysThrCysGlnGln 260
 QY 791 CTGGGTTTCAAGATGCTCAACCGGACCAACGAGGTGCCCAAGGATTTTGGCAACAGC 850
 Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGlnValAlaHisArgAspPheAlaAspSer 280
 QY 851 TTCTCAATCTTGAATCAACTCCACCATCCAGGAAAGCTTCAAGGCTTGAATGCGCT 910

Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHisArgSerGluCysPro 300
 QY 911 TCCAGAGGATATATCCCTCCAGTGTTCACCTGTCAGGCGGAGTGAAGGCGGCGG 970
 Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysSerGlyLeuArgAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGGCGGCTGGCTCGATAGCAAGTGGCTTGGCAAGTGAAGTGGTGGT 1030
 Db 321 IleValGlyIleAlaLeuAlaSerAspSerTyrTrpProIleGlnValSerLeuHisPhe 340
 QY 1031 GGCACCAACCAATCTGTGGAGGACGCTTATGACGCCCAAGTGGTGGTGGTGGTGGT 1090
 Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
 QY 1091 CACTGCTTCTTGTGACCGCGGAGAAAGTGTCTGGAGGGCTGGAAGGTGTAACGCGGAC 1150
 Db 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTyrLysValTyrAlaGlyThr 380
 QY 1151 AGCAACCTTGACACGATGCTTGAGGAGCCTCCATTCGCGAGATCATCAACAGCAAT 1210
 Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlyIleIleAsnSerAsn 400
 QY 1211 TACACCGATGAGAGAGACGACTATGACATGCGCTTCAATGCGGCTGTCCAGCCCTGACC 1270
 Db 401 TyrThrAspGluGlnAspArgTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
 QY 1271 CTGTCCGCTCATCATCCACCTGCTGGCTCCCATGATGAGACGACCTTATGACTCAAT 1330
 Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGTGATATCAGAGCCTTGGCAAGACAGGAGACAGATGACAAACATCCCCC 1390
 Db 441 GluThrCysTrpIleThrArgIlePheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
 QY 1391 TTCCTCGGGAGGTGACAGTCAATCTCATGACTTCAAGAAATGCAATGACTTGGTC 1450
 Db 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal 480
 QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGGGAGACCTTCCGTGGGCGGAGAGAC 1510
 Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyAspLeuHisGlyGlyArgAsp 500
 QY 1511 TCTGCGAGGAGAGACAGCGGGGGCTCTTGTGTGTGAGCAGAACACGCTGGTACCTG 1570
 Db 501 SerCysGlnGlnLysAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
 QY 1571 GCAAGTGTCAACAGCTGGGAGCAGAGCTGTGGCCAGAAACAAACCTGGTGTACACC 1630
 Db 521 AlaGlyValIleHisSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
 QY 1631 AAGGTGACAGAAGTCTTCCCTGGATTTACAGCAAGATGGAG 1672
 Db 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlu 554

RESULT 2

TMPSD_MOUSE STANDARD; PRT; 543 AA.
 AC Q5U405; Q8CFE0; Q91V08;
 DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
 DT 15-MAR-2005, sequence version 2.
 DT 07-MAR-2006, entry version 16.
 DE Transmembrane protease, serine 13 (BC 3.4.21.-) (Mosaic serine
 protease) (Membrane-type mosaic serine protease).
 GN Name=TMPSE13; Synonym=MeP;
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=B5/BGFP, and FVB/N;
 TISSUE=Mammary tumor, and Trophoblast stem cell;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Matusin K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tomihata S., Carninci P., Prange C.,
 RA Rha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boeck S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley D.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skelton U., Small D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
 CC protein (Potential).
 CC -1- SIMILARITY: Belongs to the peptidase S1 family.
 CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -1- SIMILARITY: Contains 1 SRCR domain.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: BC010843; AAH10843.1; -; mRNA.
 DR EMBL: BC042878; AAH42878.1; -; mRNA.
 DR EMBL: BC085323; AAH85323.1; -; mRNA.
 DR HSBP: P00760; IRX.
 DR MEROP: S01.087; -.
 DR MGI: MGI:2682935; Tmpres13.
 DR InterPro: IPR002172; LDL_rcpt_A.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_rcpt.
 DR Pfam: PF00057; Ldl_recept_a; 1.
 DR Pfam: PF00530; SRCR; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00202; SR; 1.
 DR SMART: SM00209; TYP_SPC; 1.
 DR PROSITE: PS01209; LDLR_1; FALSE_NEG.
 DR PROSITE: PS00068; LDLR_2; FALSE_NEG.
 DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
 DR PROSITE: PS0287; SRCR_2; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR GlycoProtein: Hydrolase; Membrane; Protease; Repeat; Serine protease;
 KW Signal-anchor; Transmembrane.
 FT CHAIN 1 543
 FT
 FT TOPO_DOM 1 143
 FT TRANSMEM 144 164
 FT
 FT TOPO_DOM 165 543
 FT REPEAT 14 17
 FT REPEAT 18 22
 FT REPEAT 23 27
 FT REPEAT 28 31
 FT REPEAT 32 36
 FT REPEAT 37 40
 FT REPEAT 41 45
 FT REPEAT 46 49
 FT REPEAT 50 54
 FT REPEAT 55 59
 FT REPEAT 60 64

FT REPEAT 65 69
 FT DOMAIN 180 202
 FT DOMAIN 199 301
 FT DOMAIN 302 535
 FT REGION 14 49
 FT REGION 18 69
 FT COMBIAS 11 72
 FT ACT_SITE 342 342
 FT ACT_SITE 390 390
 FT ACT_SITE 487 487
 FT CARBOHYD 231 231
 FT CARBOHYD 268 268
 FT CARBOHYD 381 381
 FT CARBOHYD 421 421
 FT DISULFID 226 290
 FT DISULFID 239 293
 FT DISULFID 327 343
 FT DISULFID 424 493
 FT DISULFID 456 472
 FT DISULFID 483 511
 FT CONFLICT 281 281
 FT CONFLICT 475 475
 FT CONFLICT 530 530
 SQ SEQUENCE 543 AA; 5906 MW; 61026D04A0FCE2D5 CRC64;
 Alignment Scores:
 Pred. No.: 4,46e-133 Length: 543
 Score: 2510.50 Matches: 468
 Percent Similarity: 90.4% Conservative: 40
 Best Local Similarity: 83.3% Mismatches: 35
 Query Match: 75.3% Indels: 19
 DB: 1 Gaps: 4
 US-10-806-370-11 (1-1748) x TMPSD_MOUSE (1-543)
 QY 11 ATGAGAGGAGGACGACCGGGAATGCTTCACGACCAAGAACACTTACGTCGACGATCT 70
 DB 1 MetAsparGlySerHisArgAsnSerProAlaArgThr----- 14
 QY 71 CCAAGCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGCCAGGATCT 130
 DB 15 ProProGlnAlaSerProAlaArgThrSerProAlaArgAla----- 28
 QY 131 CCAAGCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGCCAGGATCT 190
 DB 29 ProProGlnAlaSerProAlaArgThrProPro-----GlnAlaSer 42
 QY 191 CCAAGCTGTACACTCCAGGCGGAGCATCTCCAGGCGGAGCATCTCCAGCCAGGATCT 250
 DB 43 ProAlaArgThrProPro--GlnAlaSerProAlaArgAlaProProGlnAlaSer 61
 QY 251 CCAAGCCGAGCATCTCCAGCTGGGACACCTCCAGGCGGAGCATCTCCAGGAGGATCTCA 310
 DB 62 ProAlaArgAlaSerProAlaArgAlaProProSerArgSerSerGlyArgSerSer 81
 QY 311 TCCGCGAGGTCAGGCTCGGTGACCAACTCCCAACAGAGTGTACTTGTAAAGCAACA 370
 DB 82 SerAlaArgSerAlaSerThrThrSerSerProThrArgValTyrLeuValArgAlaThr 101
 QY 371 CCAAGTGGGGCTGTACCATCCATCTCTCTCCAGGTCAGGACCAAGCAACAGAGGCC 430
 DB 102 ProValGlyValAlaProIleArgAlaSerProAlaArgSerAlaProAlaThrArgAla 121
 QY 431 ACCAGGAGAGCCAGGATACGAGCTCCCAAGTTCAGCTGCGAGGAGGAGCAAGACAG 490
 DB 122 ThrArgGlnSerProGlyLeuSerPheProGlyPheSerTrpGlnGlnArgGln 141
 QY 491 CTACCGCTCATCGAGTGGTCT 550
 DB 142 LeuProLeuIleGlyCysValIleLeuLeuIleSerLeuValIleSerLeuLeu 161
 QY 551 TTCCAGTTCTGCGAGGCGCCACAGGATACAGGTACAGGACGAGGAGAGCTGTCTCC 610


```

QY 453 GCCGCCAAGTTCACCTGGCGGAGGCGCAGACGCTACCGCTCATCGGTGCGGC 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 MetcInProlysserProserGlyThrValCysThrSerLysThrLysValLeuLys 81
QY 513 -----TCCCTCCATTCGCTCG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 LeuThrLeuThrLeuGlyThrPheLeuValGlyAlaLeuAlaAlaGlyLeuLeuTrp 101
QY 531 TGGTTTGCTCATCATCTCTTCCAGTTTGGCAGGGCCACACAGGATCAGTTACAG 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 120
QY 591 AGCAGAGGAGAGGCTGCCACAGACCGCTGCCCTGGAGGCGGTGTGACCTGCACG 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 -----CysIleSerProSerAsnTrp---CysAspGlyValSerHisCysProG 136
QY 651 TGAAGAGTACGACGTGGGCTGGCTGAGGCTTGACTGGGACAAAGTCTCTGTTAAATCT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 LysGlyLeuAspGlyAsnArgCysValArgLeuTyrglyProAsnPheIleLeuGlnValT 156
QY 711 ACTCTGGGCTCTCCCATCATGAGCTTCCCATCTGTAGCAGACACTGGAATGACTCTACT 770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 YrSerSerHisSerLysSerTrpHisProValCysGlnAspArgTrpAsnGlyLysTyrg 176
QY 771 CAGAGAAAGCTGCCAGCAGCTGGGTTCCGAGAGTCTCACCCGACCAACAGGTTGCC 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 LysArgAlaAlaCysArgAspMetGlyTyrgLysAsnAsnPheTyrgSerSerGln---GlyI 195
QY 831 ACAGAGATTTTGCACACGCTTCTCAATCTGAGATACACTCCACG----- 877
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LevalAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 215
QY 878 --ATCCAGAAAGGCTCCACAGGTCTGAA--TGCCTTCCACAGCGGTATATCTCCCTCC 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 spIerTyrgLysLeuTyrgLysLeuTyrgHisSerAspAlaCysSerSerLysAlaValAlaSerLeuA 235
QY 933 AGTGTCTCCCATCTGGGACTGAGG-----GCCATGACCGGGCGGATCGTGGAGGGG 983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 TgCysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgGlnSerArgIleValGlyGlyG 255
QY 984 CGTGGCGCTCGGATATGACAAAGTGGCTTGGCAAGTGAAGTCACTCCGCGCACCCACA 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 LysSerAlaLeuProGlyAlaTrpProTrpGlnValSerLeuHisValGlnAsnValHisV 275
QY 1044 TCTGTGAGGAGCAGCTCATTTAGACGCCAGTGGTCTCACTGCGCGCCACTGCTTCTCG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 alCysGlyGlySerIleIleThrProGluTrpIleValThrAlaHisCysVal---- 293
QY 1104 TGACCCCGGAGAGGTCTCGAGGGC-----TGAAGGTGACCGGGCACACGCA 1154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 -----GluLysProLeuAsnAsnProTrpHisIleThrAlaPheHisIleGlyIleLeuA 311
QY 1155 ACCTGACCAAGTGGCTGAGGACGCGC-----TCCATGTCGCGAGATCATTCACAGCA 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 TgGlnSerPheMetPheTyrgLysAlaGlyTyrgLysValGlnLysValIleSerHisProA 331
QY 1209 ATTAACACGATGAGAGGAGCAGTATATGACATCGCCCTTCATCGCGCTGCAAGCCCTGA 1268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 snTyrgAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 351
QY 1269 CCGTGTCCGCTCACATCCACCTGCTTCCCTCCCATGATGACAGACCTTTAGCCCTCA 1328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetLeuLeuGlnProG 371
QY 1329 ATGAGACCTGTGTGATCAGAGCTTTGGCAGAGACAGGAGACAGATACACAGATCCC 1388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 LysGlnLeuCysTrpIleSerGlyTrpGlyAlaThrGlnGlu---LysGlyLysThrSerG 390
QY 1389 CTTTCTCCGGGAGGTGAGGTCAATCTCATGACTGACTTCACAGAAATGACATGACTTGG 1448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 LuValLeuAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgTyrgV 410

```

```

QY 1449 TCTATGACAGTACTTATACCCACAGATGATGTGTGCTGGGAGACCTTGTGGGGCAGAG 1508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 alTyrgAsnAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 430
QY 1509 ACTCTCCACAGGAGACAGCGGGGGGCTCTGTGTGTGACAGACACACCGCTGTACC 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 spSerCysGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTrpTrpL 450
QY 1569 TGGCAGGTGTCCACAGCTGGGGGCGACAGCTGTGGCCAGAGAAACAAACCTGTGTACA 1628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyrgArgProGlyValTyrg 470
QY 1629 CCAAGTACAGAAAGTCTTCCCTGAGATTTACAGCAAGATG 1669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 LysAsnValThrValPheThrAspTrpIleTyrgArgGlnMet 483

RESULT 5
Q311V3_9PRIM
ID Q311V3_9PRIM PRELIMINARY; PRT; 484 AA.
AC Q311V3_
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TMPRSS2;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Clark N.L., Swanson W.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL, DQ150501; AA282286.1; -, Genomic_DNA.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016020; C: membrane; IEA.
DR GO: GO:0008233; F: peptidase activity; IEA.
DR GO: GO:0005044; F: scavenger receptor activity; IEA.
DR GO: GO:0004252; F: serine-type endopeptidase activity; IEA.
DR GO: GO:0006508; P: proteolysis; IEA.
DR InterPro: IPR002172; LDL_rcpt_A.
DR InterPro: IPR001254; Peptidase_S1_S6.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR001190; Srcr_rcpt..
DR Pfam: PF00057; Ldl_recept_a; 1.
DR Pfam: PF00089; Tryptsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00192; LDLa; 1.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS01209; LDLa_1; 1.
DR PROSITE: PS0068; LDLa_2; 1.
DR PROSITE: PS0287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Protease; Transmembrane.
FT NON_TER 1
FT NON_TER 484
SQ SEQUENCE 484 AA; 53153 MW; 3D197DA4442BB6C5 CRC64;

Alignment Scores:
Pred. No.: 8,05e-36 Length: 484
Score: 783.50 Matches: 184
Percent Similarity: 50.9% Conservative: 69
Best Local Similarity: 37.0% Mismatches: 189
Query Match: 23.5% Indels: 55
DB: 2 Gaps: 17

```

```

US-10-806-370-11 (1-1748) x Q311V3_9PRIM (1-484)
QY 288 GGTCCATATCGGAGGATCAT---CATCCGCGAGATGACCTCGGTGACAACTCCCA 344
Db 7 G1yProtyrTylGluAenH1sG1yTfclnProGluAenPro-----TyrProAlaGln 24
QY 345 CCAGAGTGT-----ACCTGTGTAGAGCAACAGCAGTGGGGGCTGTACCCATCC----- 392
Db 25 ProthrValAlaProthrValTyrGluValH1sProAlaGlnTyrTyrProSerProVal 44
QY 393 GATCATCTCCCTCCAGGATGACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
Db 45 ProGlnTyrAlaProAlaGValLeuThrGlnAlaSerAenPro-----ValValArg 61
QY 453 GCGTGGCCAGATTACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512
Db 62 ThrGlnProtyrSerProSerGlyThrValCysThrSerTyrTyrLysValLeuGly 81
QY 513 -----TCTCTCTCATTCGCTCG 530
Db 82 11lethrLeuThrLeuGlyThrPheLeuValG1yAlaAlaLeuAlaAlaGlyLeuLeuTyr 101
QY 531 TGGTTTGGCTCATATCTCTTCCAGTTTGGCAGGGCCACACAGGGATCAGTACAAAG 590
Db 102 LysPheMetGlySerTyrCysSerApsSerGlyIleGluCysApsSerGlyThr--- 120
QY 591 AGCAGAGGAGAGAGCTGCCCAAGCAGCTGTCGCTGAGCAGGGGAGGAGGAGGAGGAG 650
Db 121 -----CysIleSerProSerAenTyr---CysApsGlyValSerH1sCysProS 136
QY 651 TGAAGAGTGAAGAGCTGGGCTGCTGAGTGTGACCTGGACAAAGTCTTGCTTAAATCT 710
Db 136 erg1yGluApsGluAenArgCysValAlaArgLeuTyrGlyProAenPheIleLeuGlnValT 156
QY 711 ACTTGGGCTCTCCCATCATGAGCTTCCCATCTGTAGACGACAACTGGAAATCTCTACT 770
Db 156 yrserserg1nArgLysSerTyrPheProValCysGlnApsApsTyrAenGluAenTyrG 176
QY 771 CAGAGAGAGCTGCGAGAGCTGCGAGTGTGACCTGACCGGACACCGGAGGAGTGGCC 830
Db 176 1yArgAlaAlaCysArgApsPheG1yTyrLysAenAenPheTyrSerSerGln---GlyI 195
QY 831 ACAAGGATTTTGGCAACAGCTTCTCAATCTTGAGATCAACTCTCCACC----- 877
Db 195 1eValApsApsSerGlySerThrSerPheMetLysLeuAenThrSerAlaG1yAenValA 215
QY 878 --ATCCAGAAAGGCTCCACAGGTTGAA---TGCCTTCCACCGGATATCTTCCCTCC 932
Db 215 sp1eTyrTyrLysLysLeuTyrH1sSerApsAlaCysSerSerTyrAlaValValSerLeuA 235
QY 933 AGTGTCCCACTGCGAGCTGAGG-----GCCATGACCGGCGGAGTGTGGAGGAGG 983
Db 235 rGysValAlaAlaCysGlyValAenLeuAenSerSerAlaGlnSerArg1leAlaG1yL 255
QY 984 CGGTGGCTCGAGTGAAGTGGCTTGGCAAGTGAAGTGTGCACTTGGCAGCAGCAGCA 1043
Db 255 alSerAlaLeuProGlyAlaTyrProTyrGlnValSerLeuH1sValGlnAenValH1sV 275
QY 1044 TCTGTGAAGGACGCTCATTTGACCCGAGTGGGTCTCATCTGCGCCGCACTGCTTTCG 1103
Db 275 alCysGlyGlySerTyrIleThrProGluTyrIleValThrAlaAlaH1sCysVal---- 293
QY 1104 TGACCCGCGAGAGAGTCTGAGAGG-----TGAAGGTGTAGCGGCGGAGCAGCA 1154
Db 294 -----GluLysProLeuAenAenProTyrH1sTyrThrAlaPheAlaGlyLe---- 309
QY 1155 ACCTGACACAG-----TTGCTGAGGAGCAGC-----TCCATTTGCGAGATCATCATCA 1202
Db 310 --LeuAArgGlnSerPheMetPheTyrGluAlaG1yTyrGlnValG1yLysValIleSerH 329
QY 1203 ACAGCATTTACACCGAGTGAAGAGGAGCAGTATGACATGCGCTTATGCGGCTGTCCAAAG 1262
Db 1203 ACAGCATTTACACCGAGTGAAGAGGAGCAGTATGACATGCGCTTATGCGGCTGTCCAAAG 1262

```

```

Db 329 1sProAenTyrApsSerTyrThrLysAenAenApsIleAlaLeuMetLysLeuGlnLysP 349
QY 1263 CCCTGACCCCTGTCGCTGCATCATCCACCTGCTTGCCTCCCATGATGAGACAGACTTTA 1322
Db 349 roLeuThrPheAenApsLeuValLysProValCysLeuProAenProGlyMetMetLeuG 369
QY 1323 GCTCTCAATGAGACCTGCTGTGATCATCAGAGCTTTGGCAAGACCGAGGAGACAGATGACA 1382
Db 369 1uProGlnGlnLeuCysTyrPheSerGlyTyrGlyAlaThrGlnGlu---LysGlySer 388
QY 1383 CATCCCTCTCTCTCCGAGAGGTGCAAGTGCATCTCATGCACTTCAAGAAATGCAATGACT 1442
Db 388 hrSerGluValAlaAenAlaAlaArgValLeuLeuIleGluThrGlnLysCysAenSerA 408
QY 1443 ACTGTGCTATGACATTAACCTTACCCCAAGATGATGTGCTGGGAGCTGTGGAGG 1502
Db 408 rGlyTyrValTyrApsAenLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyA 428
QY 1503 GCAGAGACTCTGCGCAGAGGAGACAGCGGGGCGCTTGTCTGTGACAGAAACAGCGCT 1562
Db 428 snValApsSerCysGlnGlyApsSerGlyGlyProLeuValThrLeuLysAenAenValT 448
QY 1563 GGTACCTGGCAGGCTGTACCAAGCTGGGACAGAGCTGTGGCCAGAGAAACAACCTGCTG 1622
Db 448 rPrlLeuIleGlyApsThrSerTyrPheIleSerGlyCysAlaLysAlaTyrArgProGly 468
QY 1623 TGTACACCAAGTGAAGATTCCTCCCTGATTTTACAGCAAGATG 1669
Db 468 alTyrGlyAenValThrValPheThrApsTyrPrlTyrArgGlnMet 483

RESULT 6
Q311V5_PANTR
ID Q311V5_PANTR PRELIMINARY; PRT; 484 AA.
AC Q311V5
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Transmembrane protease serine 2 (Fragment).
GN Name=TMPRSS2;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butioria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Pan.
OX NCBI_TaxId=9598;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Clark N.L., Swanson M.J.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
EMBL, DQ150499; AA82284.1; -, Genomic DNA.
GO, GO:0016021, C:Integral to membrane; IEA.
GO, GO:0016020, C:membrane; IEA.
GO, GO:0008233, F:peptidase activity; IEA.
GO, GO:0005044, F:scavenger receptor activity; IEA.
GO, GO:0004252, F:serine-type endopeptidase activity; IEA.
GO, GO:0005088, P:proteolysis; IEA.
InterPro, IPR002172; LDL_rcp_A.
InterPro, IPR001254; Peptidase_S1_s6.
InterPro, IPR001314; Peptidase_S1A.
InterPro, IPR001190; Srcr_rcpC.
Pfam, PF00057; Ldl_recept_a; 1.
Pfam, PF00089; Trypsin_1.
PRINTS, PR00722; CHYMOTRYPSIN.
SMART, SM00192; LDLa; 1.
SMART, SM00202; SR; 1.
SMART, SM00020; Tryp_Spc; 1.
PROSITE, PS01209; LDLa_1; 1.
PROSITE, PS50068; LDLa_2; 1.
PROSITE, PS50287; SRCR_2; 1.
PROSITE, PS50240; TRYPSIN_DOM; 1.

```



```

QY 288 GGTCCATCCGAGGATCAT---CATCCGAGATGAGCTCGATGACAACTCCCA 344
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 7 G1yProGlyTyrGluAenH1eG1yTyrGlnProGluAenPro-----TyrProAlaGln 24
QY 345 CCAGAGTGT-----ACCTGTTAGAGACAGACAGTGGGGGCTGTACCCATCCGATCAT 398
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 25 ProThrValAlaProAenValTyrGlnuAenH1eProAlaGlnTyrTyProSer----- 42
QY 399 CTCCTGCGAGGTCCAGACACAGACAGGCGCCAGGAGAGAGCCAGGT----- 449
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 43 -----ProValProGlnTyrThrProArgValLeuThrHis 54
QY 450 -----CGAGCTGC-----CGAAGTCACTGGCGGAGGCGCAGAGCAGC 491
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 55 AlaSerAenProAlaValCyAenGlnProLysSerProSerGlyThrValCyAenThrSer 74
QY 492 TACGGCTCATGGGGTGGTGC----- 512
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 75 LysThrIlyValAlaLeuCyAenValThrMetThrLeuGlnAlaValAlaValAla 94
QY 513 ---TCCTCCTCATTTGCCCTGGTGGTTCGTCATCATCTCTCCAGTTCTGCGAGGCGC 569
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 95 LeuAlaAlaGlyLeuLeuThrIlyPheMetGlySerIlyCySerAenSerGlyIleGln 114
QY 570 ACACAGGATCAGTACAGAGAGAGAGAGAGCTGCCAAGCAGCTGTGCTGTG 629
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 115 CyAenSerSerGlyThr-----CyA1SerSerSerAenTyr--CyA 129
QY 630 ACGGGGGTGTGATGATGACAGAGTGAAGTGAAGTGAAGTGGGTGGTGAAGTTTACTGG 689
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 129 ArgGlyValSerH1eCyAenProAenGlnIlyAenArgGlyAenAlaArgCyAenValArgLeuTyrGlyP 149
QY 690 ACAAGTCTGCTTAAATCTACTCTGGGCTGCCATCATCTGCTCCATCTGACAG 749
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 149 roAenPheIleLeuGlnValTyrSerSerGlnArgIlySerTyrH1eProAlaCyAenArg 169
QY 750 GCAACTGGATGATCTCTACTCAAGAAAGCTGCGACAGCTGGGCTTCCAGAGTCTG 809
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 169 sPaerTyrAenGlnAlaSerIlyAlaArgAlaAlaCyAenArgMetMetGlyTyrIlyAenSerP 189
QY 810 ACCGAGACACCGAG--GTTGCCACAGAGGATTTGGCAGAGCTTCCATCTTGAGAT 866
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 189 heTyrSerSerGlnIlyIleValAenAenSerGlyAlaThrSerPheMetIlyLeuAenT 209
QY 867 AC-----AATCCACATCCAGAAAGCTTCCACAGCTTGA---TGCCCTTCCC 914
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 209 hrSerAlaGlyAenValAenPheIlyTyrIlyValSerIlyH1eSerAenAlaCySerSerI 229
QY 915 AGCGGTATCTCTCCATGAGTTCCTCCATCGGAGCTGAGG-----GCCATGACCG 965
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 229 ysaValaValSerLeuAenArgSerIleAlaCyAenValArgSerAenLeuSerArgGlns 249
QY 966 GCGGATCGTGGAGGCGCGCTGCGCTCCATGACAGTGGCTTGGCAATGAGTCTGC 1025
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 249 eraGlyIleValGlyGlnAenH1eLeuLeuGlnAlaTyrProTyrProGlnAenSerIleuH 269
QY 1026 ACTTGGCAGCAGCAGCAGTGTGTGAGGAGCAGCTCATTTGACGCCAGTGGTGTCTAGT 1085
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 269 lsaValGlnAenH1eIleValCyAenGlySerIleIleThrProGlnTyrIleAlaThrA 289
QY 1086 CCGGCGCATGCTTTTGTGTGATCCCGGAGAAAGTCTCGAGAGG-----TGAAAG 1136
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 289 lalaValaCyAenVal-----GlnIlyPyrProLeuAenSerProTyrIleThrIlyA 305
QY 1137 TGATGCGGAGGAGCAGAGACCTGACAGCTGG-----CTGAGAGCAGCTCCATTCGCG 1190
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 305 lareValaGlyThrLeuArgGlnSerSerMetPheTyrGlnIlyGlyH1eArgValGln 325
QY 1191 AGATCATCATCAGACAGATTAACCCAGTGAAGAGAGAGAGATATGATGAGCTCCATGCG 1250
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 325 ysaValIleSerH1eProAenTyrAenSerIlyThrIlyAenAenAenPheIleAlaLeuMetL 345
QY 1251 GAGTGTCCAGAGCCCTGACCTGTCCGCTCCATTCACCTGCTTGGCTCCCATGATG 1310

```

```

Db 345 ysaLeuH1eThrProLeuThrPheAenGlnuValIlyPyrProValCyAenProAenProG 365
QY 1311 GACAGACCTTTAGCTCAATGAGCCTCGTGAATCATCAGAGCTTGGAGAGAGGAG 1370
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 365 lymetMetLeuGlnProGlnGlnH1eCyAenTyrIleSerGlyTyrGlyAlaThrGlnu 384
QY 1371 CAGATGACAGACATCCCTCTCCGAGAGGTGACATCATCTCATCTTCAAGA 1430
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 385 --LysGlyIlyThrSerAenPValLeuAenAlaIleMetValProLeuIleGlnProArg 404
QY 1431 AATGCAATGACTACTTGTGTATGACAGTTACTTACCCCAAGATGATGTGCTGGG 1490
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 404 rgCyAenAenH1eTyrValIlyAenPlyAenPlyLeuIleThrProAlaMetIleCyAenIleGlyP 424
QY 1491 ACCTTGGTGGGCGAGAGACTCCCTGCGAGGAGAGACAGGAGGCGCTGTGCTGTAGC 1550
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 424 heLeuGlnGlnIlyThrValAenSerCyAenGlnIlyAenSerGlyIlyPyrProLeuValThrLeu 444
QY 1551 AGAACAAACCGCTGTACTGACAGTGTCAACAGCTGGGCGACAGGCTGTGCGCAGAGA 1610
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 444 ysaAenAenPValTyrTyrLeuIleGlyAenPheThrSerTyrGlySerGlyCyAenAlaGlnAla 464
QY 1611 ACAAACTGGTGTGTACCAAGAGTGAAGAGATTTCTCCCTGGATTTACAGCAAGTG 1669
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 464 sNaArgProGlyValTyrGlyAenValThrValPheThrAenTyrIleTyrArgGlnMet 483

RESULT 9
ID 06P7D7_RAT PRELIMINARY; PRT; 490 AA.
AC 06P7D7;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 18.
DE Transmembrane protease, serine 2.
GN Name=Trmp82;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Caminici P., Mullaly S.J.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mulvaney S.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huily S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kesteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smallue D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX Strauberg R.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License

```


RESULT 10
Q3UKES3 MOUSE PRELIMINARY; PRT, 490 AA.
ID Q3UKES3
AC Q3UKES3
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 21-FEB-2006, entry version 7.
DE 14 days pregnant adult female placenta cDNA, RIKEN full-length
DE enriched library, clone:153026G04 product:transmembrane protease,
DE serine-2, full insert sequence.
GN Name=Trpsr2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Methods Enzymol. 303:19-44(1999).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX PubMed=1641072; DOI=10.1126/science.1112014;
RA Carninci P., Kankawa T., Katayama S., Gough J., Frich M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodius R., Shimokawa K.,
RA Balic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmink L.G., Adkins J., Allen J.E.,
RA Amesl-Imphibato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Barsal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chaik A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engelstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furum M., Futaki S., Gariboldi M.,
RA Georgiev-Hemmings P., Gingras T.R., Gojobori T., Green R.E.,
RA Guetlich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Itohawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Klatano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottacini-Tebber S., Mulder N., Nakano N., Nakanchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J.C., Reid J.F., Ring B.Z., Ringwald M.,
RA Rest B., Ruan Y., Salzberg S.L., Sander A., Schneider C.,
RA Schombach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA Tamajo K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yeig K.,
RA Yamashita H., Zabarovsky E., Zharov R., Zimmer A., Hilde W., Bulc C.,
RA Grimmond S.M., Teasdale R.D., Liu B.T., Brusic V., Quackenbush J.,
RA Wohlschuetz C., Wietick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki U., Arakawa T.,
RA Iida U., Imanura K., Itoh M., Kato T., Kawai H., Kawagashita N.,
RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Niimiyi N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome";
RL Science 309:1559-1563(2005).
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX PubMed=1641073; DOI=10.1126/science.1112009;
RG (Genome Network Core team) and the FANTOM Consortium;

RT "Antisense Transcription in the Mammalian Transcriptome";
RL Science 309:1564-1566(2005).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;
RA Okazaki Y., Furum M., Kankawa T., Adachi J., Bono H., Kondo S.,
RA Okazaki I., Osato N., Saito R., Suzuki H., Yamashita H., Kiyosawa H.,
RA Yeig K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bulc C., Hume D.A., Quackenbush J.,
RA Schirral L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake U.A., Bradt D., Brusic V., Chochla C., Corbani L.E., Cousins S.,
RA Dalla E., Dregani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaesteland T., Gariboldi M., Giesl C., Godzik A., Gough J.,
RA Grimmond S., Guetlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kodzler R.M., King B.L.,
RA Kankawa A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sander A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
RA Verardo R., Wagner L., Wohlschuetz C., Wang Y., Watanabe Y., Wells C.,
RA Wilmink L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayashizaki N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Itoh Y., Itoh M., Kageawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yamashita H., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Biene E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M.,
RA Arakawa T., Hara A., Pukhriehi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Iwama M., Nishi K., Kiyosawa H., Kondo S., Yamashita H.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kankawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Flieschmann W., Gaesteland T., Giesl C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirral L.M., Strubli F., Suzuki R., Tomita M., Wagner L., Watanabe T.,
RA Sakai K., Okido T., Furum M., Kono H., Baldarelli R., Baren G.,
RA Blake U., Botelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momtaz P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakomoto N.,
RA Sasaki H., Sato K., Schombach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayashizaki N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
[7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Placenta;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "IKEN Integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NOCLOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Placenta;
RA Akawa T., Carinici P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida Y., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Kono H., Murata M., Nakamura M., Niinomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watanishi A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL: AK146047; BAE26558.1; -; mRNA.
CC MGI: MGI:1354381; Tmpres2.
DR MGI: GO:0016021; C:Integral to membrane; RCA.
DR Interpro: IPR002172; LDL_recp_A.
DR Interpro: IPR001254; Peptidase_S1_S6.
DR Interpro: IPR001314; Peptidase_S1A.
DR Interpro: IPR001190; Srcr_rcpt.
DR Pfam: PF00057; Ldl_recept_a; 1.
DR Pfam: PF00089; Trypsin_1.
DR PRINTS: PR00722; CHMOTRPSIN.
DR SMART: SM00192; LDla; 1.
DR SMART: SM00020; SR; 1.
DR SMART: SM00020; Tryp_spec; 1.
DR PROSITE: PS01209; LDLRA_1; 1.
DR PROSITE: PS01209; LDLRA_2; 1.
DR PROSITE: PS50068; LDRA_2; 1.
DR PROSITE: PS50287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR Hydrolase: Membrane; Protease; Serine protease; Transmembrane;
KW SEQUENCE 490 AA; 53499 MW; 54651AF288E76EBA CRC64;
SQ
Alignment Scores:
Pred. No.: 3,82e-35 Length: 490
Score: 771.50 Matches: 174
Percent Similarity: 47.3% Conservative: 87
Best Local Similarity: 31.5% Mismatches: 200
Query Match: 23.1% Indels: 91
DB: 2 Gaps: 13
US-10-806-370-11 (1-1748) x Q3UKE3_MOUSE (1-490)
QY 89 GCTGGGACACCTCCAGGCGGGGATCTCCAGCCAGGATCTCCAGCCAGGATCTCCA 148
DB :::::::::::::::::::::
QY 5 Ser1ySerProProGly---1LeGlyProCyStrglnuSnH1sglyTyrglnSerGlu 23
DB :::::::::::::::::::::
QY 149 GCTGGACACCTCCGGGGCGGGGATCTCCAGCCAGGATCTCCAGCTGGTACATCCA 208
DB :::::::::::::::::::::
QY 24 His1leCySPProPro-----ArgProPro 31
DB :::::::::::::::::::::
QY 209 GAGCGGACATCTCAGGCGGGGATCTCCAGCCAGGATCTCCAGCC-----CGG 259
DB :::::::::::::::::::::
QY 32 ValAlaProAenglyTyraSnLeuTyProAlaGlnTyTyProSerProAlaProGln 51
DB :::::::::::::::::::::
QY 260 GCATCTCGGCTTCGATCATCTTCCAGGCTCCATCCGATCCGCGAGATCATCCGCCAG 319
DB :::::::::::::::::::::
QY 52 TyrlaProAlaGlyleThnThrglnAlaSerThrservAl1leH1stnH1sProLysSer 71
DB :::::::::::::::::::::
QY 320 TCAGCCTCGGTGACAACTCCCAACCAAGATGATACCTTTGTTAGACAAACCAAGTGGG 379
DB :::::::::::::::::::::

DB 72 SerGlyAlaLeuCyStrHsrSerLySerTylybSerLeuCybLeuAlaLeuAlaLeuGly 91
QY 380 GCTGTACCATTCGATCATCTCTCCAGAGTCAAGCAACCAAGGCGCACAGGAG 439
DB :::::::::::::::::::::
QY 92 ThrVal----- 93
DB :::::::::::::::::::::
QY 440 AGCCCAAGTACAGAGCTGCCCAAGTTCACTGGCGGGAGGCCAAGACAGTACCGCTC 499
DB :::::::::::::::::::::
QY 93 ----- 93
DB :::::::::::::::::::::
QY 500 ATGGGAGTGGCTGCTCTCCATATGCGCTGGTGTTCCTCATATCTCTTCCAGTTC 559
DB :::::::::::::::::::::
QY 94 -----LeuThrglyAlaAlaValAlaAlaValAlaLeuLeuTyPrArgPhe 107
DB :::::::::::::::::::::
QY 560 TGGCAGGCGCCACACAGGATTCAGGTTCAGAGCAGAGAGAGC-----TGTCCCAAG 613
DB :::::::::::::::::::::
QY 108 TyraSPserSnHnYserThrservGlnMeSerGlyCybGlySerSerGlyThrcyAl1Ser 127
DB :::::::::::::::::::::
QY 614 CAGCTGTGGCTGTGACGGGGGTGGAGTGGACSTGAAAGTGAAGTGAAGTGGGCTGC 673
DB :::::::::::::::::::::
QY 128 SerSerLeuThrcySerArgGlyValAlaH1sCybProHsnGlyGlnArgGlnAlaArgCyb 147
DB :::::::::::::::::::::
QY 674 GTAGAGTTTACCTGGCAAGTCTGCTTTAAATTAATTAATTAATTAATTAATTAATTA 733
DB :::::::::::::::::::::
QY 148 ValArgLeuTyrglnGlnSerPhe1leLeuGlnValTySerSerGlnArgLyAlaTrp 167
DB :::::::::::::::::::::
QY 734 CTTCGCACTCTGTACAGCAACATGTAAGTCTCTACATCAGAGAAAGTCTGCGACAGCTG 793
DB :::::::::::::::::::::
QY 168 TyrProValCybGlnaPaaPrrSerGlnSerTyrlaArgAlaAlaCybLybAspMet 187
DB :::::::::::::::::::::
QY 794 GGTTCGAGAGTGTCAACCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 850
DB :::::::::::::::::::::
QY 188 G1yTyrlaSnHn 207
DB :::::::::::::::::::::
QY 851 TTTCAATCTTGAGATCAATCTC-----ACATTCAGAAAGCTTCCACAG 898
DB :::::::::::::::::::::
QY 208 PheMetLybLeuHnValSerSerGlyAsnValAspLeuTyrlaLybLeuTyrlaH1Ser 227
DB :::::::::::::::::::::
QY 899 TCTGAATGCCCTTCCAGCGGATATATCTCTCCAGTTCCTCCAGTCCGAGTGGAGGGCG 958
DB :::::::::::::::::::::
QY 228 AspSerCybSerSerArgMetValValSerLeuArgGlySerGlyValAlaH1Ser 247
DB :::::::::::::::::::::
QY 959 ATG-----ACGGGCGGATCTGTTGGGGGCGGCTGCTGATGATGATGATGATGATGATG 1012
DB :::::::::::::::::::::
QY 248 ValLybArgGlnSerArgGlyValGlyGlyLeuHnHnHnHnHnHnHnHnHnHnHnHnHnHnHnHn 267
DB :::::::::::::::::::::
QY 1013 CAAGTGAATGTGCACTTCGACACCAACCAATCTGTGAGAGCAGCTATTCAGCCAG 1072
DB :::::::::::::::::::::
QY 268 GlnValSerLeuHn 287
DB :::::::::::::::::::::
QY 1073 TGGGTGCTCACTGCGCGGCACTGCTTTGTGTGACCCGGGAGAAAGTCTTGGAGGCG--- 1129
DB :::::::::::::::::::::
QY 288 Trp1leValThnHn 303
DB :::::::::::::::::::::
QY 1130 -----TGAAGGTGAGCGGCGACCC-----AGCAACTG 1159
DB :::::::::::::::::::::
QY 304 ArgTyrlaThnHn 323
DB :::::::::::::::::::::
QY 1160 CACCAAGTTCGCTGAGGAGGCTTCATTCGAGATCAATCAACAGCAATTCACCGAT 1219
DB :::::::::::::::::::::
QY 324 HisGln-----ValGlnValVal1leSerH1sProHnTyraSPser 337
DB :::::::::::::::::::::
QY 1220 GAGGAGACAGACTATGATTCGCTTCATGCGGTTCGCAAGCCCTGACCTTCGCT 1279
DB :::::::::::::::::::::
QY 338 LysThnLybAsnHn 357
DB :::::::::::::::::::::
QY 1280 CACATCCACCTGCTGCTCTCCCATCATGACAGACACTTTAGCTCAATGACAGCTGCG 1339
DB :::::::::::::::::::::
QY 358 LeuValLybProValCybLeuProSerProGlyMetLeuLeuArgGlnGlyCyb 377
DB :::::::::::::::::::::
QY 1340 TGGATCAAGGCTTTGGCAAGACAGGAGACAGATACAGACATGCTCCCTTCTCGCG 1399
DB :::::::::::::::::::::
QY 378 Trp1leSerGlyTrpGlyAlaThnTyrgln---LysGlybTybThrservAlaValLeuHn 396
DB :::::::::::::::::::::


```

OY 734 CTTCCCATCTGTAGCAGCACTGGAAATGACTCTTACTCAGAGAAAGCCTGCGCAGAGCTG 793
DB 168 TyrProValCysGlnAspAspTrpSerGlnSerGlnValGlnAlaCysValAspMet 187
OY 794 GGTTCGAGAGTGTCTACCGGACCAACCGAG--GTTGCCACAGGATTTTGGCAACAGC 850
DB 188 GlnTyrLysAsnAsnMetLysSerSerGlnGlnLysLeuProAspGlnSerGlnAlaThrSer 207
OY 851 TTCTCAATCTTGAGATACACTCC-----ACCATCCAGAAAGCCTCCACAGC 898
DB 208 PheMetLysLeuAsnValSerSerGlnValAsnValAspLeuTyrLysLeuTyrHisSer 227
OY 899 TCTGAATGCCCTTCCAGCGGATATATCTCCCTCCAGTTTCCCACTGGGACGTAGGCGCC 958
DB 228 AspSerCysSerSerArgMetValValSerLeuValGlnGlnCysGlnValAlaArgSer 247
OY 959 ATG-----ACCGGCGGATGCTGGGAGAGGCGCTGCGCTGCGATGACAAAGTGGCCTTGG 1012
DB 248 ValLysArgGlnSerArgGlnValGlnValGlnLeuAsnAlaSerProGlnValAspTrpTrp 267
OY 1013 CAACTGACTGTGCACTTGGGACCAACCACTGTGTGGAGGACGCTCATTGACGCCAG 1072
DB 268 GlnValSerLeuHisValGlnGlnValHisValCysGlnGlnSerLeuLeuThrProGln 287
OY 1073 TGGGTGCTCACTGCGCGCCCACTGCTTCTTCTGTCAGCCGCGAGAGAGTCTTGAGAGGC--- 1129
DB 288 TrpLeuValThrAlaAlaHisCysVal-----GlnGlnProLeuSerSerPro 303
OY 1130 -----TGAAGAGTGTACGCGGCGCC-----AGCAACTG 1159
DB 304 ArgTyrTrpThrAlaPheAlaGlnGlnLeuArgGlnSerLeuMetPheTyrGlnSerArg 323
OY 1160 CACCACTGCTGAGGACGCTCCATTCGAGATCATCATCAACAGCAATTATACCCGAT 1219
DB 324 HisGln-----ValGlnLysValHisSerHisProAlaThrAspSer 337
OY 1220 GAGAGAGCACTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTTACCCCTGCGCT 1279
DB 338 LysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnThrProLeuAlaPheAsnAsp 357
OY 1280 CACATCCACCGCTGCTGCTCCCATGATGATGAGACCTTATGCTCATGAGACCTGC 1339
DB 358 LeuValLysProValCysLeuProAsnProGlnMetLeuAspLeuAspGlnGlnCys 377
OY 1340 TGGATCACAGGCTTGGCAAGACGAGGACAGATGACAGACATCCCTTCCCTCCG 1399
DB 378 TrpLysSerGlnTyrGlnAlaThrTyrGln---LysGlnLysThrSerAspValLeuAsn 396
OY 1400 GAGGTGACGTCATCTCATGACTTCAAGAAATGCAATGACTACTTGGTCTATACAGT 1459
DB 397 AlaAlaMetValProLeuIleGlnProSerLysCysAsnSerLysTyrLysAsnAsn 416
OY 1460 TACCTTACCCCAAGATGATGTCGCGGACCTTCCGCGGCGGACGACCTCCGCGAG 1519
DB 417 LeuIleThrProAlaMetIleCysAlaGlnPheLeuGlnGlnSerValAspSerCysGln 436
OY 1520 GGAGACAGCGGGGCGCTTGTCTGTGAGACAAACCGCTGCTGCTGCGCAGTCTC 1579
DB 437 GlnAspSerGlnGlnProLeuValThrLeuLysAsnGlnLysLeuTrpTrpLeuIleGlnAsp 456
OY 1580 ACCAGCTGGGACAGGCTGTGGCCAGAAACAACCTGCTGTGTATCAACAAAGTACA 1639
DB 457 ThrSerTrpGlnSerGlnCysAlaLysAlaLeuAspProGlnValTyrGlnValAsnValThr 476
OY 1640 GAAGTCTTCCCTGATTATACAGAGATGAGAGC 1675
DB 477 ValPheThrAspTrpIleTyrGlnMetAlaGln 488

```

```

DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE Transmembrane protease, serine 2 variant.
GN Name=TPRSS2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.; to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA MEDLINE=94171032; PubMed=8125298; DOI=10.1016/S0378-1119(94)90802-8;
RX "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=96038986; PubMed=9373149; DOI=10.1016/S0378-1119(97)00411-3;
RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama K., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RL end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Tanaka A., Yokoyama S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL: BC051839; AAHS1839.1; -; mRNA.
DR EMBL: AK222784; BAD96504.1; -; mRNA.
DR Ensembl: ENSG00000184012; Homo sapiens.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016020; C: membrane; IEA.
DR GO: GO:0008233; F: peptidase activity; IEA.
DR GO: GO:0005044; F: scavenger receptor activity; IEA.
DR GO: GO:0004252; F: serine-type endopeptidase activity; IEA.

```

DR GQ:0006508; P:proteolysis; IEA.
 DR InterPro: IPR002172; LDL_rcpt_A.
 DR InterPro: IPR01254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Srcr_rcpt.
 DR Pfam: PF00089; Trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLa_1.
 DR SMART: SM00202; SR_1.
 DR SMART: SM00202; Tryp_spec_1.
 DR PROSITE: PS01209; LDLa_1; 1.
 DR PROSITE: PS50068; LDLa_2; 1.
 DR PROSITE: PS50287; SRCR_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase: Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 492 AA; 53859 MW; C05B531C8A311C7 CRC64;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
5.64e-35	768.50	492	178
Percent Similarity:	50.64	Conservative:	74
Best Local Similarity:	35.74	Mismatch:	195
Query Match:	23.14	Indels:	51
		Gaps:	15

US-10-806-370-11 (1-1748) x Q6GTK7_HUMAN (1-492)

QY 288 GGTCTCTCATCCGGCAGGTGAT---CATCCGCCAGGTGACCTCGGTGACAACTCCCA 344
 Db 12 G1yProTyrTyrGluasnHisLeGlyTyrGlnProGluasnPro-----TyrProAlaGln 29
 QY 345 CCACAGAGT-----ACCTGTGAGCAGCAGCAGTGGGGCTGTACCTCC----- 392
 Db 30 ProThrValValProThrValTyrGlnValHisProAlaGlnTyrTyrProSerProVal 49
 QY 393 GATCATCTCTCCGACGATGACACACACAGCAGGCGCCACGAGGAGAGCCAGCATGCA 452
 Db 50 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValValCys 66
 QY 453 GCCGCCCAAGTTCACCTGGGGAGGAGGCCAGAAAGCAGCTACCGCTCATCGGGTGGCTGC 512
 Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysAlaLeuCys 86
 QY 513 -----TCTCTCTCATGCGCTGG 530
 Db 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaLeuAlaAlaGlyLeuLeuTrp 106
 QY 531 TGGTTTGGCTCATCTCTCTTCAGAGTTCGACAGGCGCACACAGGATCAGTACAG 590
 Db 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
 QY 591 AGCAGAGGAGAGAGCTGCCCAAGCAGCTGCTCGTGTGACGGGGTGTGACATGCAAGC 650
 Db 126 -----CysIleAsnProSerAsnTrp--CysAspGlyValSerHisCysProG 141
 QY 651 TGAAGAGTGAAGAGCTGGCTGGCTGAGAGTTTGAATGAGCAAGCTCTGCTTAAATCT 710
 Db 141 LysGlyGluAspGluAsnArgCysValAlaGlyLeuTyrGlyProAsnPheIleLeuGlnValT 161
 QY 711 ACTCTGGAGTCTCCCATCATGAGCTTCCCATCTGTAGACAGCACTGAAATGACTTACT 770
 Db 161 YrSerSerGlnArgLysSerTrpHisProValCysGlnAspAspTrpAsnGluAsnTyrG 181
 QY 771 CAGAGAAAGCTGCCAGAGCTGGGTTTCAGAGTGTCTCAGCCGACCAACAGAGTTGCC 830
 Db 181 LysAlaAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
 QY 831 ACAGAGATTGTCACCAAGCTTTCATCTGAGAGTACACTCCAGC----- 877
 Db 200 LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220

QY 878 --ATCCAGAAAGCTCCACAGGTCTGAA---TCCCTTCCACAGCGGTATATCTCCCTCC 932
 Db 220 spIeTyrLysLysLeuThrHisSerAspAlaCysSerSerValAlaValIleSerLeuA 240
 QY 933 AGTTTCCCATCTGGGACTGAG-----GCCATGACCCGGGCGGATCGTGGAGGG 983
 Db 240 rGcysIleAlaCysGlyValAsnLeuAsnSerSerArgGlnSerArgIleValIdylGlyG 260
 QY 984 CGCTGGCCCTGGATAGCAAGTGGCTTGGCAAGTGTGCACTTGGCGGACCAACCA 1043
 Db 260 IuSerAlaLeuProGlyAlaTrpProTrpGlnValSerLeuHisValGlnAsnValHisV 280
 QY 1044 TCTGTGAGAGCAGCCTCATTTGACCCCAAGTGGTGTCTCATCTGCCGCCACTGTTCTTCG 1103
 Db 280 alCysGlyLysSerIleIleThrProGluTrpIleValAlaThrAlaAlaIleCysVal---- 298
 QY 1104 TGACCCGGAGAAAGTCTCTGAGGGC-----TGAAGGTGTAGCGGGGACCAACA 1154
 Db 299 -----GluLysProLeuAsnAsnProTrpHisTrpThrAlaPheAlaGlyIleLeuA 316
 QY 1155 ACTGCACACAGTTCCTGAGCAGCAGC-----TCCATTTGCCAGATCATCTCAACAGA 1208
 Db 316 rGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGluLysValIleSerHisProA 336
 QY 1209 ATTAACACCGATGAGAGAGCAGTATGACATCGCCTTCATGCGGCTGTCCAGCCCTGA 1268
 Db 336 snTyrAspSerLysThrLysAlaAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
 QY 1269 CCGTGTCCGCTCATCATCCCTCGCTTGGTCCCATGATGACACACACCTTTAGCCTCA 1328
 Db 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetLeuGlnProG 376
 QY 1329 ATGACACCTGCTGATCATCCCAAGATGATGTGTGTGGGACCTTCGTTGGGGCAGAG 1388
 Db 376 IuGlnLeuCyStrPrlSerGlyTyrPrlValAlaThrGluGlu---LysGlyLysThrSerG 395
 QY 1389 CTTCTCTCCGGAGAGTCAAGTCAATTCATTCGACTTCAAGAAATGCAATGACTACTTGG 1448
 Db 395 IuValLeuAsnAlaAlaIleValLeuLeuIleGluThrGlnArgCysAsnSerArgTyrV 415
 QY 1449 TCTATGACAGTACCTTATCCCAAGATGATGTGTGTGGGACCTTCGTTGGGGCAGAG 1508
 Db 415 alTyrAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435
 QY 1509 ACTCTGCGCAGAGAGACAGCGGGGGCTCTTGTGTGTGACAGAAACAACGCTGTATCC 1568
 Db 435 epSerCysGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnIleTrpTrpL 455
 QY 1569 TGGCAGGTGTACCAAGCTGGGGCAGCAGCTGTGGCCAGAGAAACAACCTGTGTATCA 1628
 Db 455 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyrArgProGlyValTyrG 475
 QY 1629 CCAAGAGACAGAAAGTCTTCCCTCGATTTACAGCAAGATGGAAGGAG 1678
 Db 475 LysAsnValMetValPheThrAspTrpIleTyrArgGlnMetArgAlaAsp 491

RESULT 13
 Q31UB_HYLSY PRELIMINARY; PRT; 484 AA.
 AC Q31UB;
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 1.
 DE Transmembrane protease serine 2 (Fragment).
 GN Name=TPRSS2;
 OS Homo sapiens (Human) (Homo sapiens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Hylobactidae; Symphalangus.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Clark N.L., Swanson W.J.;


```

Db      67 ThrGlnProIySerProSerGlyThrValCysThrSerIyThrIyLysAlaLeuCys 86
QY      513 -----TCCTCTCATTCGCCCTGG 530
Db      87 IleThrLeuThrIleuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuIleuTrp 106
QY      531 TGGTGGCTGCATCATCTCTTCAGATTCCTGGACAGGGCCACACAGAGGATGATGACAG 590
Db      107 LysPheMetGlySerIyLysCysSerAsnSerGlyIleGlyCysAspSerSerGlyThr--- 125
QY      591 AGCAGAGGAGAGAGCTGTCCCAAGCAGCTGTTCGCTGTGACGGGGGTGTGAGTCAAGC 650
Db      126 -----CysIleAsnProSerAsnTrp--CysAspGlyValSerHisCysProG 141
QY      651 TGAAGAGTGAAGAGCTGGGCTGCTGCTGAGCTTGAAGTTCGAGCAAGTCTCTGTTAAATCT 710
Db      141 LysGlyIuAspGluAsnAspGlyValAlaGlyLeuTrpGlyProAsnPheIleLeuInMet 161
QY      711 ACTCTGGGCTCTCCATCATGAGCTGCTCCATCTGTAGACAGACATGGAATGACTCTACT 770
Db      161 YrSerSerGlnArgIyLysSerTrpHisProValCysGlnAspAspTrpAsnGluAsnTrpG 181
QY      771 CAGAGAAGACCTGCGACAGAGCTGGGTTTCAGAGTGTCTACCGGACCAACCGAGTGGCC 830
Db      181 LysArgAlaAlaCysArgAspMetGlyTrpLysAsnAsnPheTyrSerSerGln---GlyI 200
QY      831 ACAGGAGATTGGCCAAAGCTTTCATCTTGAATACAACTCCACC-----877
Db      200 LeValAspAspSerGlySerThrSerPheMetCysLeuAsnThrSerAlaGlyAsnValA 220
QY      878 --ATCCAGGAAGGCTCCACAGAGCTGAA--TGCCCTTCCACGGGATATCTCCCTCC 932
Db      220 spIleTrpLysLysLeuTrpHisSerThrAspAlaCysSerSerLysAlaValSerLeuA 240
QY      933 AGTGTTCCTCCATGCGGAGCTGAGG-----GCCATGACCGGCGGATCTGGAGAGGG 983
Db      240 rgCysIleAlaCysGlyValAsnLeuAsnSerSerArgGlnSerArgIleValGlyGlyG 260
QY      984 CGTGGGCTCGGATGACAGAGTGGCTTGGCAAGTGTCTGACTTCGGACCAACCCACA 1043
Db      260 IuSerAlaLeuProGlyValAlaTrpProIleGlnValSerLeuHisValGlnAsnValHisV 280
QY      1044 TCTGTGAGGAGGACCTTATTAACCCGAGTGGGTGTCTACGCGGCCACATGCTTCTCG 1103
Db      280 alCysGlyGlySerIleThrProGluTrpIleValThrAlaAlaHisCysVal--- 298
QY      1104 TGACCCCGGAGAGAGTCTGTGAGGAGC-----TGGAAGTGTACCGCGGACACAGCA 1154
Db      299 -----GluLysProLeuAsnAsnProTrpHisTrpThrAlaPheAlaGlyIleLeuA 316
QY      1155 ACCTGCACCACTGCTCGAGGACGCC-----TTCATTGCCAGATCATCAACAGCA 1208
Db      316 rgGlnSerPheMetPheTrpGlyAlaGlyTrpGlnValGlnValIleSerHisProA 336
QY      1209 ATTACACCGAGAGAGAGACATATGACATGCGCTCATGCGCTGCTCCAGGCCCTGA 1268
Db      336 enIyAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
QY      1269 CCTGTGCGCTGCATATCCACCTGCTGCTCCCATGATGACAGACCTTTAGCTCA 1328
Db      356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetCysLeuGlnProG 376
QY      1329 ATGAGACCTGTGTGATCAAGAGCTTGGCAAGACCGAGGAGACAGATGACAAAGATCC 1388
Db      376 IuGlnLeuCysTrpIleSerGlyTrpGlyAlaThrGluGln---LysGlyLysThrSerG 395
QY      1389 CCTTCTCCGAGAGAGTGCATGATCATGATGATGATGATGATGATGATGATGATGATG 1448
Db      395 IuValIleAsnAlaAlaLysValLeuLeuIleGluThrGlnArgCysAsnSerArgIyrv 415
QY      1449 TCTATGACAGTTACTTACCCCAAGGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1508
Db      415 alTyraAspAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435

```

```

QY      1509 ACTCTGCGCAGGAGAGACAGCGGGGCTCTGTGTGTGACGACAAACCGCTGTAC 1568
Db      435 spSerCysGlnGlyAspSerGlyValProLeuValThrSerLysAsnAsnIleTrpIPL 455
QY      1569 TGGCAGGTGTACACAGCTGGGGGACAGCTGTGGCCAGAGAAACAACTGTGTGTACA 1628
Db      455 euIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyraArgProGlyValTyrg 475
QY      1629 CCAAAGTGACAGAAAGTTCTTCCTGATTTTACAGCAAGATGAGAGCGCAG 1678
Db      475 LysAsnValMetValPheThrAspTrpIleTyraArgGlnMetArgAlaAsp 491

RESULT 15
ID      096T73_HUMAN      PRELIMINARY;      PRT;      492 AA.
AC      096T73;
DT      01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT      01-DEC-2001, sequence version 1.
DE      21-FEB-2006, entry version 26.
DE      Epithelialasin.
GN      Name=TPRS32.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC      Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NCUBLOTIDE SEQUENCE.
RX      MEDLINE=21223025; PubMed=1132890;
RA      Jacquinet E., Rao N.V., Rao G.V., Wang Z., Albertine K.H.,
RA      Hoidal J.R.;
RT      "Cloning and characterization of the cDNA and gene for human
RT      epithelialasin."
RL      Eur. J. Biochem. 268:2687-2699(2001).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL: AF329454; AAK53559.1; -; mRNA.
DR      HSSP; P00766; 1CHG.
DR      Ensembl; ENSG00000184012; Homo sapiens.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005044; F:scavenger receptor activity; IEA.
DR      GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR      GO; GO:0006508; P:proteolysis; IEA.
DR      InterPro; IPR002172; IDL_rcpt_A.
DR      InterPro; IPR001254; Peptidase_S1_S6.
DR      InterPro; IPR001314; Peptidase_S1A.
DR      InterPro; IPR001190; Src_rcpt.
DR      Pfam; PF00089; Trypsin; 1.
DR      PRINTS; PRO0722; CHYMOTRYPSIN.
DR      SMART; SM00192; IDLA; 1.
DR      SMART; SM00202; SR; 1.
DR      SMART; SM00020; TRYR_SPC; 1.
DR      PROSITE; PS01209; IDLA_1; 1.
DR      PROSITE; PS50068; IDLA_2; 1.
DR      PROSITE; PS50287; SRCR_2; 1.
DR      PROSITE; PS50240; TRYPsin_DOM; 1.
DR      PROSITE; PS00134; TRYPsin_His; UNKNOWN_1.
DR      PROSITE; PS00135; TRYPsin_SER; 1.
KW      Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
SQ      SEQUENCE 492 AA; 53863 MW; 3ABA755B276DADF CRC64;

Alignment Scores:
Pred. No.:      7,31e-35      Length:      492
Score:      766.50      Matches:      178
Percent Similarity:      50.6%      Conservative:      74
Best Local Similarity:      35.7%      Mismatches:      195
Query Match:      23.0%      Indels:      51
DB:      Gaps:      15
US-10-806-370-11 (1-1748) x 096T73_HUMAN (1-492)

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 01:52:32 : Search time 12.6 Seconds
(without alignments)
3642.941 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 ctcagagaccatgagagaggg.....ggctctgctgactcgagaaa 1748

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 650591 seqs, 87530628 residues
Total number of hits satisfying chosen parameters: 1301182

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DEV=xlp
-Q=abseq/ABSSWEB.spool/US10806370/runat.15092006.105634.10230/app.query.fasta_1
-DB=Issued_Patents_AA -OPMT=fastaan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -INITS=b1ts -START=1 -END=-1 -MATRIX=blomun62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abseq02p -USER=US10806370 @CGN_1_1_84 @runat.15092006.105634.10230
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /EMC_Celerra_SIDS3/pcodata/2/1aa/5_COMB.pep:*
2: /EMC_Celerra_SIDS3/pcodata/2/1aa/6_COMB.pep:*
3: /EMC_Celerra_SIDS3/pcodata/2/1aa/7_COMB.pep:*
4: /EMC_Celerra_SIDS3/pcodata/2/1aa/H_COMB.pep:*
5: /EMC_Celerra_SIDS3/pcodata/2/1aa/ECTUS_COMB.pep:*
6: /EMC_Celerra_SIDS3/pcodata/2/1aa/RE_COMB.pep:*
7: /EMC_Celerra_SIDS3/pcodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES
Result No. Score Query Match Length DB ID Description

1	2999	60.8	562	2	US-09-879-792-12	Sequence 12, Appl
2	2291.5	68.8	477	2	US-10-177-661-2	Sequence 2, Appl
3	2108	63.2	446	2	US-10-177-661-4	Sequence 4, Appl
4	969.5	29.1	418	2	US-10-177-661-6	Sequence 6, Appl
5	769.5	23.1	492	2	US-09-759-143-932	Sequence 932, App
6	769.5	23.1	492	2	US-10-012-896-932	Sequence 932, App
7	769.5	23.1	492	3	US-10-144-678A-932	Sequence 932, App
8	768.5	23.1	492	2	US-09-342-749-2	Sequence 2, Appl
9	768.5	23.1	492	2	US-09-691-840-2	Sequence 2, Appl
10	768.5	23.1	510	2	US-09-949-016-11074	Sequence 11074, A
11	764.5	22.9	492	2	US-09-685-166A-895	Sequence 895, App

12	764.5	22.9	492	2	US-09-879-792-14	Sequence 14, Appl
13	764.5	22.9	492	2	US-09-679-426-895	Sequence 895, App
14	764.5	22.9	492	2	US-09-759-143-895	Sequence 895, App
15	764.5	22.9	492	2	US-10-012-896-895	Sequence 895, App
16	764.5	22.9	492	3	US-10-144-678A-895	Sequence 934, App
17	764.5	22.9	492	3	US-09-759-143-934	Sequence 934, App
18	736.5	22.1	393	3	US-10-012-896-934	Sequence 934, App
19	736.5	22.1	393	3	US-10-144-678A-934	Sequence 934, App
20	736.5	22.1	521	2	US-09-949-016-11081	Sequence 11081, A
21	736.5	22.1	521	2	US-09-949-016-11082	Sequence 11082, A
22	736.5	22.1	521	2	US-09-949-016-11083	Sequence 11083, A
23	729.5	21.9	453	2	US-09-999-833A-69	Sequence 69, Appl
24	729.5	21.9	453	2	US-10-020-445A-69	Sequence 69, Appl
25	729.5	21.9	453	2	US-09-978-189-69	Sequence 69, Appl
26	729.5	21.9	453	2	US-10-017-085A-69	Sequence 69, Appl
27	729.5	21.9	453	3	US-10-145-129A-69	Sequence 69, Appl
28	729.5	21.9	453	3	US-10-013-929A-69	Sequence 69, Appl
29	729.5	21.9	453	3	US-10-013-917A-69	Sequence 69, Appl
30	726	21.8	454	2	US-09-518-046-2	Sequence 2, Appl
31	726	21.8	454	2	US-09-650-371-2	Sequence 2, Appl
32	704	21.1	423	2	US-09-656-002-2	Sequence 2, Appl
33	697.5	20.9	435	2	US-09-607-743-2	Sequence 2, Appl
34	697.5	20.9	435	3	US-10-030-688-2	Sequence 2, Appl
35	697.5	20.9	437	2	US-09-851-588-8	Sequence 8, Appl
36	696.5	20.9	435	2	US-09-008-271A-6	Sequence 6, Appl
37	696.5	20.9	435	2	US-09-968-415-6	Sequence 6, Appl
38	695.5	20.9	406	2	US-09-851-588-6	Sequence 6, Appl
39	695	20.9	492	3	US-10-030-688-4	Sequence 4, Appl
40	694	20.8	432	2	US-10-012-231A-275	Sequence 275, App
41	694	20.8	432	2	US-10-015-389A-275	Sequence 275, App
42	694	20.8	432	2	US-10-006-768A-275	Sequence 275, App
43	694	20.8	432	2	US-10-015-671A-275	Sequence 275, App
44	694	20.8	432	2	US-10-015-393A-275	Sequence 275, App
45	694	20.8	432	2	US-10-011-833A-275	Sequence 275, App

ALIGNMENTS

RESULT 1
US-09-879-792-12
: Sequence 12, Application US/09879792
: Patent No. 6734006
: GENERAL INFORMATION:
: APPLICANT: Xiao, Yonghong
: APPLICANT: Gedrich, Richard
: TITLE OF INVENTION: Regulation of Human Transmembrane Serine
: FILE REFERENCE: 02973.00035
: CURRENT APPLICATION NUMBER: US/09/879,792
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/211,224
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: US 60/283,353
: PRIOR FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: US 60/283,648
: PRIOR FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: PCT
: PRIOR FILING DATE: 2001-06-12
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12
: LENGTH: 562
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-879-792-12

Alignment Scores:
Pred. No.: 2,11e-218
Score: 2999.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 90.0%
DB: 2
Length: 562
Matches: 562
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

DB: 2 Gaps: 1
US-10-806-370-11 (1-1748) x US-10-177-661-2 (1-477)
QY 368 ACAACAGTGGGGCTGTACCCATCCATCTCTTCGACAGGTGACGACCAAGCAGG 427
DB 38 SerProthLeuSerIleSerSerAspSerSerProIaIaIaIa 57
QY 428 GCCACCAAGGAGAGCCAGGTACGAGCTCCCAAGTTCACCTGGCGGAGGGCCAGAG 487
DB 58 LeuLeuLeu-----GlyThrSerLeuProIySerThrItpaGgIuGInLyS 74
QY 488 CAGTACCGGTCAATGGGGTGGTGGTCTCTCTCATTCCTGGGAGTGGTTCATCATC 547
DB 75 GInLeuProIeulIeGlyCySValIeLeuLeuIleAlaLeuValIaSerIleIle 94
QY 548 CTCTTCAGTTCGTGACAGGAGCCACACAGGATCAGGTACAGAGACAGAGAGAGCTGT 607
DB 95 LeuPheGInPheItrpGInGlyIhIsthrGlyIleArgGlyIeGInuGInuSerCys 114
QY 608 CCCAAGCAGCCTGTCTGCTGTGACGGGGTGGTGGTACGTCAAGTGAAGAGTGAAGCTG 667
DB 115 ProlYhIaIaIaValArgCySAspGlyValIaAspCySlySerLeuYsSerAspGInLeu 134
QY 668 GAGTGGGTGAGTGTGACGTGGGCAAGCTCTGTGTTAAATCTACTGGGCTCTCCCAT 727
DB 135 GlyCySValArgPheAspItrpAspIlySerLeuLeuYsIleIySerGlySerSerHis 154
QY 728 CAGTGGCTTCCTCATCTGTACAGACAACTGGATGACTCTTACTCAGAGAACTGGCAG 787
DB 155 GInItrpLeuProIleCySerSerSerItrpAspAspSerIySerGInuYsthrCySgIn 174
QY 788 CAGCTGGGTTCCAGAGTGTCAACCGGACAAACGAGTGGTCCCAAGGATTTGGCAAC 847
DB 175 GInLeuGlyPheGInSerAlaHisArgThrThrGInuValaIaHisArgAspPheAlaAsn 194
QY 848 AGCTTCATCTTGATGATCAACTCCACCATCCAGAAAGCCTCCACAGTCTGAATGC 907
DB 195 SerPheSerIleLeuArgIyAsnSerThrIleGInuIleSerLeuHisArgSerGInuCyS 214
QY 908 CCTTCCAGCGGTATATCTCCCTCAAGTGTTCACATGCGAGTGAAGGCGCATGACCGG 967
DB 215 ProSerGInuArgIyIleSerLeuGInuCySerHisIeGlySerIyLeuArgAlaMetThrGly 234
QY 968 CGGATCGTGGAGGAGGCGCTGGCTCGATATGACAGTGGCTTGGCAAGTGAAGTGCAC 1027
DB 235 ArgIleValGlyGlyAlaLeuAlaSerAspSerIyItrpItrpGInuValSerLeuHis 254
QY 1028 TTGGGACCAACCAATCTGTGAGGAGCGCTCATTTAGCGCCAGTGGGTCACATGCC 1087
DB 255 PheGlyThrThrHisIleCySgIyGlyThrLeuIleAspIaGInItrpValLeuThAla 274
QY 1088 GCCCAGTCTTCTTCTGACCCGAGAGAAAGTCTGGAGGGCTGGAAGTGTAGCGCGG 1147
DB 275 AlaHisCySHeheheValThrArgGInuValValLeuGInuIyItrpYsValIyAlaGly 294
QY 1148 ACCAGCACTGTGACCAAGTGGCTGAGGAGCTTCATTCGCGAATCATCATCAAGC 1207
DB 295 ThrSerAsnLeuHisGInLeuProGInuAlaIaSerIleAlaIleIleIleAsnSer 314
QY 1208 AATTACACCGATGAGAGGAGCACTATGACATGCGCTCATGGGGCTGTCAGAGCCCTG 1267
DB 315 AsnIyThrAspGInuGInuAspAspIyAspIleAlaLeuMetArgLeuSerIySProLeu 334
QY 1268 ACCGTGCTCGCTCATCATCACTGCTTGGCTCCCATGATGAGACAGCTTTAGCTTC 1327
DB 335 ThrLeuSerAlaHisIleHisIleProIaCySLeuProMetHisGlyGInItrpPheSerLeu 354
QY 1328 AATGAGACCTGCTGGATCAAGAGCTTTGGCAAGACCAAGGAGACAGATGACAGATCC 1387
DB 355 AsnGInuThrCySArgItrpIleItrpGlyPheGlyIyAsThrArgGInuItrpAspIySthrSer 374
QY 1388 CCTTCTCTCGGAGGTGAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG 1447

DB 375 ProPheLeuArgGInuValGInuValaLeuLeuIleAspPheIySerCySAsnAspIyTrLeu 394
QY 1448 GTCTATGACAGTTACTTACCCCAAGATGATGTGTGCTGGGACCTTGTGGGGCAGAG 1507
DB 395 ValIyAspSerIyTrLeuThrProArgMetCySAlaGlyAspLeuArgGlyGlyArg 414
QY 1508 GACTCTGCGAGGAGACAGCGGGGGCCTCTGTCTGTGAGAGAAACAGCTGGTGTAC 1567
DB 415 AspSerCySgInuIyAspSerGlyGlyProLeuValCySgInuGInuAsnAspIyTrIy 434
QY 1568 CTGGCAGTGTCAACGACTGGGACAGAGCTGTGGCCAGAGAAACAGCTGTGTGTAC 1627
DB 435 LeuAlaGlyValIhIserItrpGlyThrGlyCySgIyGInuArgAsnIySProGlyValIy 454
QY 1628 ACCAAGTGAACAGAGTTCCTCTGATTTACAGCAAGTGAAGACGAGTGCATTC 1687
DB 455 ThrIyValThrGInuValLeuProItrpIleIySerIySerMetGInuSerGInuValArgPhe 474
QY 1688 AGAAATCC 1696
DB 475 ArgIySer 477
RESULT 3
US-10-177-661-4
; Sequence 4, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Varca, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299,606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-661-4
Alignment Scores:
Pred. No.: 4,29e-151 Length: 446
Score: 2108.00 Matches: 393
Percent Similarity: 89.6% Conservative: 4
Best Local Similarity: 88.7% Mismatches: 12
Query Match: 63.2% Indels: 34
DB: 2 Gaps: 2
US-10-806-370-11 (1-1748) x US-10-177-661-4 (1-466)
QY 368 ACAACAGTGGGGCTGTACCCATCCATCTCTTCGACAGGTGACGACCAAGCAGG 427
DB 38 SerProthLeuSerIleSerSerAspSerSerProIaIaIaIa 57
QY 428 GCCACCAAGGAGAGCCAGGTACGAGCTCCCAAGTTCACCTGGCGGAGGGCCAGAG 487
DB 58 LeuLeuLeu-----GlyThrSerLeuProIySerThrItpaGgIuGInLyS 74
QY 488 CAGTACCGGTCAATGGGGTGGTGGTCTCTCTCATTCCTGGGAGTGGTTCATCATC 547
DB 75 GInLeuProIeulIeGlyCySValIeLeuLeuIleAlaLeuValIaSerIleIle 94
QY 548 CTCTTCAGTTCGTGACAGGAGCCACACAGGATCAGGTACAGAGACAGAGAGAGCTGT 607
DB 95 LeuPheGInPheItrpGInGlyIhIsthrGlyIleArgGlyIeGInuGInuSerCys 114
QY 608 CCCAAGCAGCCTGTCTGCTGTGACGGGGTGGTGGTACGTCAAGTGAAGAGTGAAGCTG 667
DB 115 ProlYhIaIaIaValArgCySAspGlyValIaAspCySlySerLeuYsSerAspGInLeu 134

QY	668	GGCTGCGGAGATTGATGATGGGAAAGCTGTGGCTTAAATCTATCTCGGGTCCCAT	727
Db	135	GIYGVTAIAIgrheleBPTTPAaplyseIeueuylsIeIySeGlySeSerhIs	154
QY	728	CAGTGGCTTCCATCTGTAGGAGCACTGAATGACTCTACTAGAGAAAGCTGGCAG	787
Db	155	GIIntrPeuproIIeCyseSeSerAenItrpAmApseryIySeGluIyThCyG6In	174
QY	788	CAGCTGGGTTTCGAGAGTGTCTCACCGGACAAACGAGTTGGCCACAGGGAATTTTGGCAAC	847
Db	175	GIIneuGIyPheGlu-----	179
QY	848	AGCTTCTCAATCTTGAGATATACAACTCCAGCATCCAGGAAAGCTTCACAGTCTGAATGC	907
Db	180	-----ArgSeGluCys	183
QY	908	CCTTCCCAAGCGATATATCTCCCTCCAGTGTTCCTCCACTGCGGACTGAGGGCCATACCGGG	967
Db	184	ProSeGInAIGyTIISeSerIeGInCysSehIscyG6IyLeuAIGraIaIeThreGly	203
QY	968	CGGATCGGGAGGAGGGCGCTGAGCTCGGATATGACAAAGTGGCCTTGAGCAAGTAGTGTGAC	1027
Db	204	ArgIIeValIGyGlyAlaIeuaIaSeIzApseryIyStrProItrpGInValSeIueuIs	223
QY	1028	TTTGCGACCAACCCACATCTGTGAGAGCAGGCTCATTTGAACGCCAAGTGGTCTCACTGCC	1087
Db	224	PheGIyThrThnIstIeCySeGIy6IyThrIeueIleAraIaGIInItrpValIeueThraIa	243
QY	1088	GGCCAGTGGCTCTTCGTGACCCGGGAGAAAGTCCGTGAGGGGCTGAAAGGTATAGCGGGC	1147
Db	244	AlaIscyPheIreValThraArgGluIySValIeueGluGlyItrpIyValIyAlaGly	263
QY	1148	ACGAGCAACCTGCACCAAGTTCCTGAGGAGCAGCTCCATTTGCCAGATCATCATCAACAGC	1207
Db	264	ThreSeAenIeuehIaGIInIeuproGluhIaIaSeIeIaIaGIuIleIleIleAenSer	283
QY	1208	AATTAACACCGATGAGAGAGAGACTATGACTATCGGCTCATGCGGCTGTCCAAAGCCCTG	1267
Db	284	AenIyIThraPepGluGluAAspIyIzApsIleAlaIeueMeIzAglSeIySeIyProIeue	303
QY	1268	ACCCTGTCGGGTCAATCAACCTCGTCTGCTCCCGCATGCAATGGACACCTTTAGCTTC	1327
Db	304	ThreueSerIaIaIstIeIhIeIspToIaCySeIeupromethIaGIyGIInThPheSerIeue	323
QY	1328	AATGAGACTGTGTGATCACAGGCTTTGGCAAGACGAGAGACAGATGACAAAGATCC	1387
Db	324	AenGIuThrcyStrpIIeThrGIyPheGIyIyStrhIarGIuThraAspIyIyThSer	343
QY	1388	CCCTTCTCCGGGAGGTGCAAGTCAATCTCATGCACTTCAAGAAATGCAATGACTACTTG	1447
Db	344	ProPheueuAIGyGluValGInValaIaenIeueIleApsPheIySlyCyAAsAAspIyIreue	363
QY	1448	GTCTATGACAGTTACTCTTACCCCAAGATGATGTGTGTGGGAGCTTCGTGGGGGAGCA	1507
Db	364	ValIyIzApseryIyIeueThrProIaIghetIeIcySaIaGIyApsIeuehIarGIyGIyArg	383
QY	1508	GACTTCCTGCCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGAACACCGCTGTGAC	1567
Db	384	AspSerCySgInGIyApsSerGIyGIyProIeueValCySgInGIuInaIaenAraIyItrPyr	403
QY	1568	CTGGCAGGTTGCACAGCTGGGGGACAGGCTGTGGCCGAGAAACAAACCTGGTGTGAC	1627
Db	404	LeuAIGIyValIhIreSerItrpGIyThrGIyCySgIyInaIaGAmIySproGIyValIyIyr	423
QY	1628	ACCAAGTGCAGAAATTCTTCCCTGTGATTTTACAGCAAGATGAGAGGCGAGGTGCATTC	1687
Db	424	ThryIyValIhIrgIuValIeueProItrpIIeIySeIySmetGIuSerGIuValIyArghe	443
QY	1688	AGAAATTC	1696
Db	444	ArgIySer	446

```

RESULT 4
US-10-177-661-6
; Sequence 6, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Vitca, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299,606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(418)
; OTHER INFORMATION: Xaa = unknown
US-10-177-661-6

Alignment Scores:
Pred. No.: 4,7e-65 Length: 418
Score: 969.50 Matches: 208
Percent Similarity: 56.8% Conservative: 27
Best Local Similarity: 50.2% Mismatches: 130
Query Match: 29.1% Indels: 49
DB: Gaps: 12

US-10-806-370-11 (1-1748) x US-10-177-661-6 (1-418)
QY 602 AGCTGTCACAGACGCGTTCGCTGTGACGCGGAGTGAGACTCGACGTCAGAGAGTGAC 661
DB 1 SerCyseIlePro***Ser***TrrCyahspIylVal***AapCysePro***GlyGluasp 20
QY 662 GAGCTGGGCGTGC-----GTGAGGTTTGACTGG 688
DB 21 Glu*****CysVal*****GlyPro*****ValArgLeuTyrGly 40
QY 689 GACAAAGTCCTGCTTAATCTACTCTGGGTCCTCCATCAGTGGCTTCCATCTGTAGC 748
DB 41 Asp*****LeuLeuGlnValTyrSerSerSer*****Trr**ProValCyseSer 60
QY 749 AGCAACTGAGTATGACTCCTACTCTCAGAGAACTGCGCAGACAGTGGATTGGAGAGTCT 808
DB 61 AspartTrpahnGlnSerTyrSer*****AlaCys***GlnMetGly*****SerIle 80
QY 809 CACCGGACAAACCGAG-----GTTGCCACAGGAGATTGGCAACAGCTTCTCAATCTTG 862
DB 81 **Tyr***SerGlu***Gly*****Arg***GlyAlaAenSerPhe***LysLeu 100
QY 863 AGA-----TGCACCTCCACATCCAGGAAGC 889
DB 101 AsnValSerPro***AsnLeuLeu*****Aap***TyrTrnSer**IleGln***** 120
QY 890 CTCACAGAGGTGTGA-----TGCCCTTCCAGCGGATATCTCCTCCAGTGTTC 940
DB 121 Leu***ArgSerSerSer*****CysProSerGly***ValValSerLeuGlnCysSer 140
QY 941 -----CACTGCGGACTGAG-----GCCATGACCGGCGGATCTGTGGAGGG 982
DB 141 **GlnAspCysGlyValArgLeuAsnAla*****MetTrnSerArgIleValGlyGly 160
QY 983 GCGCTGGCTCTCGGATAGCAAGTGGCTTTGGGAAGTGAAGTGCACCTTGGCAGC---ACC 1033
DB 161 *****AlaSer***Gly***TrrProTrrGlnValSerLeuGln*****GlyVal 180

```


Qy	1040	CACATCTGTGGAGGACAGCTTCATTATGACGCCAGGGGGCTCATCGCCGCCCATCGCTTC	1099
Db	181	HisLeuCyseGlyGlySerLeuIle***Pro***TrrValLeuTrnAlaAlaHisCysVal	200
Qy	1100	TTTCGTGACCCGGGGAAGGATCTCTGGAGGGCTGGAAGGTGTACGGCGGC-----ACCAGC	1153
Db	201	***GlyArg*****LysProLeu***GlyTrp***ValPheAlaGlyIleLeuTrn***	220
Qy	1154	AACCTGCACCACTGTTCCCTGAGCGAGCC-----TCATTGCCAGATATCATCAACAGC	1207
Db	221	SerLeuHis*****Pro***Ala*****Arg***ValGlyIleIleIleIleHisPro	240
Qy	1208	AATTAC-----ACGATGAGGAGACGACTATGACATCGCCCTCATG	1248
Db	241	AsnTrp*****Ser*****LysAspAsnAspIleLeuLeuMet	260
Qy	1250	CGGCTGTCCAAAGCCCTTCAGCCCTGTCGCGCTCATATCCACCCCTGCTTGCGTCCCATGCAT	1309
Db	261	LysLeuSerLysPProLeuTrnPhe***AspTrpIleGlnProValCysLeuProAsnPro	280
Qy	1310	GGAACGACCTTTAGCCTTCATATGAGACCTGCTGATCACAGGCTTTGGCAAGACAGGAG	1365
Db	281	GlyGln***Leu***ProGlyTrnThrCysTrpIle***GlyTrpGlyAlaThr***Glu	300
Qy	1370	ACAGATACACAGACATCCCCCTCTCTCCGGAGGGAGTGCAGTCACTTCATGACTTCAG	1422
Db	301	*****GlyLysTrnSerProValLeuGlnIle***ValProLeuIleAspAsnLys	320
Qy	1430	AAATGCATATGACTCTTGGTGTATGACAGTTAACCTTACCCCAAGATGATGTCCTGGG	1485
Db	321	***CysAsnSerTrp***ValTrpAspAsn***IleTrnProAsnMetIleCysAlaGly	340
Qy	1490	GACCTTGCTGGGGGCGAGACACTCTGCCAGGAGACAGCGGGGCTTTGTCGTGAG	1545
Db	341	TyrLeuGlnGlyGlyValAlaPseCysGlnGlnLysPseRgGlyProLeuValCysGln	360
Qy	1550	-----CAGAACACCGCTGGTACTCTGGCAGGTGTCCACAGCTGGGGGACAGGC	1599
Db	361	*****GlnAsnAsnArgTrpTrpLeu***Gly***ThrSerTrpGly***Gly	380
Qy	1598	TGT---GGCCAGAGAAACCAACCTGGGTGTGACCAAGTGACAGAACTTCTCCCTGG	1654
Db	381	Cys***AlaLysAlaAsnLysProGlyValTrp***ValTrn***PheLeu***Trp	400
Qy	1655	ATTATCAGCAAGATGAGCGAGCGAGGTGGATTCAGAAATATCC	1696
Db	401	IleTyrSerGlnMet***AlaGln***ArgPheArgLysSer	414
RESULT 5			
US-09-759-143-932			
; Sequence 932, Application US/09759143			
; Patent No. 6800746			
; GENERAL INFORMATION:			
; APPLICANT: Xu, Jiangchun			
; APPLICANT: Dillon, Davin C.			
; APPLICANT: Mitcham, Jennifer L.			
; APPLICANT: Harlocker, Susan L.			
; APPLICANT: Jiang, Yugui			
; APPLICANT: Henderson, Robert A.			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Fanger, Gary R.			
; APPLICANT: Retter, Marc W.			
; APPLICANT: Stolk, John A.			
; APPLICANT: Day, Craig H.			
; APPLICANT: Vedvick, Thomas S.			
; APPLICANT: Carter, Derrick			
; APPLICANT: Li, Samuel			
; APPLICANT: Wang, Aijun			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Hepler, William			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND			
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER			
; FILE REFERENCE: 210121.427C23			

```

CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PR1
ORGANISM: Homo sapiens
US-09-759-143-932

Alignment Scores:
Pred. No.: 6,47e-50 Length: 492
Score: 769.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
Db: 2 Gaps: 15

US-10-806-370-11 (1-1746) x US-09-759-143-932 (1-492)
QY 288 GGTCTTCATCCGGCAGAGTCAT--CATCCCGCAGGTACGCTCGGTACCAACTCCCA 344
Db 12 G1PProTyrTyrGluAenH1sg1YTrG1nPProG1uAnPPro-----TyrPProAlaGln 29
QY 345 CCAGAGAGT-----ACCTGTTAGACACACACAGGCGGGCGTGAACCCATCC----- 392
Db 30 ProThrValValProThrValTyrGlnValH1sProAlaGlnTyrTyrProSerProVal 49
QY 393 GATCATCTCTCCGCGAGGTCCAGACACCAACCAAGGCGCCACGAGGAGAGCCAGTACGA 452
Db 50 ProGlnTyrAlaProAlaGValLeuThrGlnAlaSerAnPro-----ValValCys 66
QY 453 GCCTGCCCAAGTTACCTGTGGCGGAGGCCAAGACAGCTACCGCTATCGGTCGCTGC 512
Db 67 ThrGlnProLysSerProSerG1YThrValCysThrSer1YsThrLysAlaLeuCys 86
QY 513 -----TCCTCCTCATTCGCCCTGG 530
Db 87 1leThrLeuThrLeuG1YThrPheLeuValG1YAla1aLeuAla1aG1YLeuLeuTrp 106
QY 531 TGGTTTGGTCATCATCTCTCTCCAGTTCTGGACGGCCACACAGGATCAGTACGAAG 590
Db 107 LysPheMetG1YSerLysCysSerAsnSerG1Y1leG1uCySAspSerSerG1YThr 125
QY 591 AGCAGAGGAGAGCGCTGCCAAGACGACGCTGTTGCTGTGACGGGTGGTGACTGCAAC 650
Db 126 -----Cys1leAsnProSerAsnTrp--CysAspG1YValSer1YsCysProg 141
QY 651 TGAAGAGTGAAGAGCTGGGCTGCTGTGAGGTTTACCTGGAGACAACTCTGTTAAATCT 710
Db 141 1Yg1Yg1uAspG1uAsnArG1YsValArG1YLeuTyrG1YSerAsnPhs1leLeuG1nValT 161
QY 711 ACTCTGGAGTCTCCCATCAGAGCGTTCGCCATCTGTATAGACAGCACTGGAATATCTACT 770
Db 161 YrSerSerGlnArG1YsSerTrpH1sProValCysGlnAspAspTrpAsnG1uAnTyrG 181
QY 771 CAGAGAAGACCTGCAGCAGCTGGGTTTCAGAGTGTCTACCGGACAAACGAGTTGCC 830
Db 181 1YarG1a1a1aCysArG1sAspMetG1Yr1YrLysAsnAsnPhetYrSerSerGln--GlyI 200
QY 831 ACAGAGATTGTTCCCAACAGCTTTCATCTTGAGATACAACTCCACC----- 877
Db 200 1eVal1aAspAspSerG1YserThrSerPheMetLysLeuAsnThsSer1aG1YAsnVala 220
QY 878 --ATTCAGGAAGAGCTCCACAGGCTGGA--TGCCCTTCCACAGCGGATATATCCCTCC 932
Db 220 sp1leTyrLysLysLeuTyrH1sSerAspAlaCysSerSer1YsAlaValValSerLeuA 240
QY 933 AGTGTTCCTCACTGGGAGTGAAG-----GCCATGACCGGCGGATGCTGGAGAGGG 983
Db 240 rGcYs1leAlaCysG1YValAsnLeuAnSerSerArG1nSerArG1leValG1Yg1Yg 260
QY 984 CGCTGGCCTGGATAGCAAGTGGCTTGGCAAGTGAATCTGCACTTGGCACCAACCA 1043

```

```

Db      260  luserAlaLeuProGlyAlaTTrpProTrpGlnValSerLeuHisValGlnAsnValHisv 280
      1044  TCTGTGAGGACGACCTATGACCCGAGTGGCTGCTACATGCCGCCACGACTCTTCTCG 1103
      280  aTCyGgLygLySerLlelleThrProGluTrpLleValThnAlaHisCysVal---- 298
      1104  TGACCCGGAGAAAGTCTGTGAAGGC-----TGAAGGTGTACGGGGACCCAGCA 1154
      299  -----GluTyProLeuAsnAsnProTrpHisTrpThrAlaPheLleGlyLleLeuA 316
      1155  ACCTGACACCACTTCCCTGAGGACCC-----TCCATTGCCGAATCATCTCAACGCA 1208
      316  rglInsErPheMetPheTyGlyAlaGlyTyGlnValGluTyValLleSerHisProA 336
      1209  ATTACACCGATGAGAGAGACATATGACATGCGGCTGCTGATCGGCTGCTGCAAGCCCTGA 1268
      336  enTyraPseTrLysThyLeuAsnAsnLleAlaLeuMetLysLeuGlnLysProLeuT 356
      1269  CCTGTCCGCTGACATCCACCTGCTGCTGCCCTCCCATGATGACAGACGCTTATGCTCA 1328
      356  hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
      1329  ATGAGACCTGTGATTCACAGGCTTTGGCAAGACCGAGAGACAGATGACAGACATCCC 1388
      376  lueGlnLeuCyStrpLleSerGlyTyTrpGlyAlaThrGluGlu---LysGlyLysThrSerG 395
      1389  CCTTCCCTCCGGAGAGTGCATATCATCTGACTTCAAGAAATGCAATGACTACTCTTG 1448
      395  luValLleAsnAlaLysValLleuLeuLleGluThrGlnArgCysAsnSerArgTyrv 415
      1449  TCTATGACAGTACTTACCCCAAGATGATGTGTGTGGGACCTTGTGGGGGACAG 1508
      415  aTyraPseLeuLleThrProAlaMetLleCysAlaGlyPheLeuGlnGlyAsnValA 435
      1509  ACTCTCTCCAGGAGACAGCGGGGGCTCTTGTCTGTGAGCAAAACAACCGCTGTGTACC 1568
      435  spSerCyGgGlnGlyAspSerGlyGlyProLeuValThrSerLysAsnAsnLleTrpTrpL 455
      1569  TGGAGAGTGTCAACAGCTGGGGACAGGCTGTGGCGGAAACAAACCTGTGTGTACA 1628
      455  euLleGlyAspTrpHisTrpLysSerGlyCysAlaLysAlaTyraTrpProGlyValTyrg 475
      1629  CCAAGTGTACAGAAATTCTTCCCTGATTTTACAGCAAGATGAGAGCGAG 1678
      475  lyaEnValMetValPheThrAspTrpLleTyraTrpGlnMetArgAlaAsp 491

```

```

APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-932

Alignment Scores:
Pred. NO.: 6,47e-50 Length: 492
Score: 769.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: 2 Gaps: 15

US-10-806-370-11 (1-1748) x US-10-012-896-932 (1-492)
      288  GGTCTCATCCGCGAGGTCAT---CATCCCGCAGTCAAGCTCGGTGACAACTCCCAA 344
      12  GLyProTyTrpGlnAsnHisGlyTyGlnProGlnAsnPro-----TyProAlaGln 29
      345  CCAGAGTGT-----ACCTGTTAGACAAACCCAGTGGGGGCTGTACCATCC----- 392
      30  ProThrValValProThrValTyGlnValHisProAlaGlnTyTrpProSerProVal 49
      393  GATCATCTCTCCGACAGTCAAGCAACCAAGGCGCACAGGAGAGAGCCAGGTACGA 452
      50  ProGlnTyraLysProAspGlyLleuThrGlnHisAsnPro-----ValValCys 66
      453  GCTGTCCCAAGTTCACCTGTGGGGAGGCGCAAGCAGTACCGCTCATCGGTGCGTGC 512
      67  ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysValAlaLeuCys 86
      513  -----TCCTCTCATTTGCCCTGG 530
      87  lLeThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaGlyLeuLeuTrp 106
      531  TGGTTTGCATCATCATCTCTCCAGTTCTGGGAGGCGCACACAGGATCAGGTACAG 590
      107  LysPheMetGlySerLysCysSerAsnSerGlyLleGlyCysAspSerSerGlyThr--- 125
      591  AGCAGAGGAGAGAGCTGTCCCAAGCACCGCTGTGCTGTGACGGGGTGTGAGTGAAGC 650
      126  -----CysLleAsnProSerAsnTrp---CysAspGlyValSerHisCysProG 141
      651  TGAAGAGTGAACAGCTGGCTGGCTGAGTTTGTGACTGGGACAGTCTGTCTTAAATCT 710
      141  LygLyGlnAspGlnAsnArgCysValArgLeuTyrgLysSerAsnPheLleLeuGlnValT 161
      711  ACTGTGCTCTCCATCAGTGGCTTCCCATCTGTGATGATGATGATGATGATGATGAT 770
      161  TySerSerGlnAspGlySerTrpHisProValCysGlnAspAspTrpHisGlnAsnTyrg 181
      771  CAGAGAGACCTGTGACAGCTGTGTTTGAAGTGTGACCGGACCAACGAGGTTGCC 830
      181  LysArgAlaAlaCysArgAspMetGlyTyraLysAsnAsnAsnPheTySerSerGln---GlyI 200
      831  ACAGGATTTTCCCAACAGCTTCTCAATCTTGAGATCAACTCCACC----- 877
      200  LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
      878  --ATCCAGAGAAAGCTCAGAGCTGGA---TGCCTTCCCAAGCGGTATATATCCCTCC 932
      220  spLleTyLysLysLeuTyHisSerAsnAlaCysSerSerLysValValValSerLeuA 240

```

```

QY 933 AGTGTCCCACTGCGGAGCTGAGG-----GCCATGACGGGGGAGTGTGGAGGG 983
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 rgcysilealaCyvglyValaslnleuanserSerarglnserArgileValgly 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 984 CGCTGCGCTCGGATAGCAAGTGGCTTGGCAAGTGAATGTGCACTTGGGACCA 1043
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 luseralaleuProglYAlatrProtrpoglInvalSerleuHleValGlnanValHisv 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1044 TCTGTGGAGCGAGCTGATGAGCCCGAGTGGGTGCTCACTGCGCCCACTGCTTTCG 1103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 alCyvglySerlelelthrProglutrpIleValIthralaahleCyvaja---- 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1104 TGACCCGGGAGGAGTCTGAGGGC-----TGAAGGTATGACCGGACCA 1154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 -----GluIyPProleuansnansProtrpIstPrtrpIalAphaaglylleuA 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1155 ACCTGACCAAGTTGCTGAGGACGC-----TCCATTGCGGAGATCATCAACAGCA 1208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 rglInserPheMetPheTyrGlyAlaGlyTyrGlnValGluIyValIleSerHisProA 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1209 ATTACACCGATGAGAGAGAGACATAGATCGCCCTCATGCGGCTGTCCAGCCCTGA 1268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 entYrAspserLysThrLysAsnAsnspIleAlaIleuMetLysLeuGlnIlyProleut 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1269 CCCTGCTCCGCTACATCCACCTGCTGCTGCTCCCATGATGACAGACCTTAACTCA 1328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 hrPheAsnAspLeuValIlyAspValCyIleuProAsnProglYMetMetLeuGlnProG 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1329 ATGAGACCTGTGATGAGAGGCTTGGCAAGACAGAGAGACAGATGACAAACATCCC 1388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 luginleuCystrIleSerlelYrPolYAlatrGluGlu--LysGlyLysHisSerG 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1389 CCTTCTCCGGAGAGTGAAGTCAATCTCATCACTTCAAGAAATGCAATGACTACTGG 1448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 luvallleuansnAlaAlaIyValleuIleGluIthrGlnArgCyAsnSerargYrv 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1449 TCTATGACAGTACTTACCCCAAGATGATGTGTGCTGGGAGCTTCGTGGGGAGAG 1508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 alYrAspAsnLeuIleThrProAlaMetIleCyAsnIleGlyPheLeuGlnIlyAsnValA 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1509 ACTCTGCGGAGGAGAGAGGCGGCGCTGTGCTGAGACAGAAACCGCGGTACC 1568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 sPserCyGlnGlyAspSerlelYrProleuValThrSerLysAsnAsnIletrpIrl 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1569 TGGAGAGTGTACCACTGCGGAGAGAGCTGTGCGCAGAGAAACAACTGTGTGTACA 1628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 eulIleGlyAspThrSerItrpIlelYserGlyCyAsnAlaIyAlaIyArgProglYAlaIyrg 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1629 CCAAGTGAAGAAATTCTTCTGATTTACAGCAAGATGAGAGCGAG 1678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 lYasnValMetValPheThrAspTrpIleYrArgIleMetArgAlaAsp 491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Baasolds, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Matanabe, Yoshitiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: DIAGNOSIS OF PROSTATE CANCER
CURRENT APPLICATION NUMBER: US/10/144,678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ. ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 932
LENGTH: 492
TYPE: PR
ORGANISM: Homo sapiens
US-10-144-678A-932
Alignment Scores:
Pred. No.: 6,47e-50 Length: 492
Score: 769.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: Gaps: 15
US-10-806-370-11 (1-1748) x US-10-144-678A-932 (1-492)
QY 288 GGTTCATCCGCGAGGTACT---CATCCCGAGTCACTGCTGATCAACTCCCA 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 GlyProIyYrGluIyAsnHisGlyIyGlnProGluAsnPro-----TyrProAlaGln 29
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 CCAAGAGT-----ACCTGTAGAGCAACACAGTGGGGGCTGTACCATCC----- 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 ProThrValValProThrValIyGlnValIleHleProAlaIleIyYrProserProval 49
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 GATCATCTCTGCGAGTCAAGACCAACCAAGGCGCACAGGAGAGAGCCAGGTACGA 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50 ProGlnIyAlaProAlaArgValleuThrGlnAlaSeranPro-----ValValCys 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 453 GCTGCCCAAGTTCACCTGCGGAGGCGGAGAGCAAGCAAGCTTACCGCTATCGGTGCG 512
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ThrGlnProIySerProserGlyThrValCyThrSerIyThrIyIyValIleuCyS 86
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 -----TCCTCTCATTTGCCCTGG 530
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 lIeThrIeuthrIeuglyThrPheLeuValGlyAlaIleuAlaIleuAlaGlyLeuIytrp 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 TGGTTTCGTCATCATCTCTTCAGTTCTGAGGAGGCGCACAGAGATCAGGTACAGG 590
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 lYsPheMetGlySerIyLysCySerAsnSerGlyIleGluCyAspSerSerGlyThr--- 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 AGCAGAGAGAGAGCTGTCCCAAGACAGCGCTGTGCTGTGACGGGTGTGACTGCAAGC 650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 -----CyAlaIeAsnProSerAsnIytrp--CyAspGlyValIleSerHisCyAsp 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 TGAAGAGTGAAGAGCTGGGCGGTGAGTTTACCTGCGACAAAGTCTGTGTTAAATCT 710
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 lYelGlyIleAspGlnAsnArgCyValIyArgLeuIyYrIySeranPheIleuGlnValI 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 711 ACTCTGGGTCTCCCATCACTGCTTCCATCTGTAGACCAACTGGAATGACTCTACT 770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 ytrSerSerGlnArgIySerItrpIstProvalCyGlnAspAspIyAsnGlnuAsnIyrg 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 771 CAGAGAAAGCTGCGACAGACTGGGTTTCAAGAGTGTCAACCGGACAAACGAGTTGCC 830
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 lYArgAlaAlaCyAsnAspMetCylYrIyAsnAsnAsnPheYrSerSerGln--GlyI 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 831 ACAGGATTTTGCAACAGCTTCTCAATCTTGAGATCAACTCCACC----- 877

```

```

Db      200  leValaAspSerGlySerThrSerPheMetLeuValSerAlaGlyLeuValA 220
      878  --ATCCAGAAAGCCCTCAGAGTCTGAA---TGCCCTTCACCGCGATATCTCCCTCC 932
Db      220  splleTyLysLeuValSerThrSerPheAlaCysSerSerLeuValValSerLeuA 240
      933  AGTGTTCACCTGCGAGCTGAGG-----GCCATGACCGCGCGATCTGGAGAGGG 983
Db      240  rgCysAlaCysGlyAlaLeuValLeuValSerSerArgGlnSerArgIleValGlyGly 260
      984  CGCTGCGCTCCGATAGCAAGTGGCTTGGCAAGTGAAGTGTGACTTCCGACCCACCA 1043
Db      260  luseRAlaLeuProGlyAlaTrpProTyrGlnValSerLeuHleValGlnValHleSv 280
      1044  TCTGTGAGGACGACCTTGAACGCCGAGTGGAGCTACGCGCGCCACGCTTCTTCG 1103
Db      280  alCysGlyGlySerIleIleThrProGlnTrpIleValThrAlaAlaHleCysVal--- 298
      1104  TGACCCGGAGAGAGTCTCTGAGAGGC-----TGAGAGTGTACCGCGACACAGCA 1154
Db      299  -----GluTyrProLeuValLeuValSerProTyrPheIleThrAlaPheAlaGlyLeu 316
      1155  ACCTGACCAAGTGGCTGAGGACGCC-----TCCATTGCCAGATCATCATCAACAGCA 1208
Db      316  rgGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGlnValIleSerHleProA 336
      1209  ATTACACCGATGAGAGAGACATATGACATGCGCTTCATGCGCGCTGCTCCAGGCCCTCA 1268
Db      336  snTyArpSerLeuThrLysAsnAsnArgIleAlaLeuMetLeuLeuGlnLysProLeuT 356
      1269  CCTGTGCGCTCAATCCACCTGCTGCTCCCTCCCATGATGACAGACACTTTCAGCTTCA 1328
Db      356  hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnPro 376
      1329  ATGAGACCTGCTGATCATCAGAGCTTTGGCAAGACCAGGAGACAGATGACAAAGCATCC 1388
Db      376  lGlnLeuCysTrpIleSerGlyTrpGlyAlaThrGlnGlu--LysGlyLysThrSerG 395
      1389  CCTTCCTCCGGAGAGTGGAGTGCATATCTCATGCACTTCAGAAATGCAATATCATCTTG 1448
Db      395  luvAlleuValAlaAlaValLeuValLeuValLeuGlnThrGlnArgCysAsnSerArgTyr 415
      1449  TCTATGACATTAACCTTATACCCAGAGATGATGTGTGCTGGAGGAGCTTGGAGGAGAG 1508
Db      415  alTyArpAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnLysValAla 435
      1509  ACTCTTCGAGGAGACAGACCGGAGGCTTGTCTGTGACAGACAGAACCGCTGTGATCC 1568
Db      435  spSerCysGlnGlyAspSerGlyLysProLeuValThrSerLysAsnValIleTrpTrl 455
      1569  TGGAGAGTGTACACAGCTGGGAGGACAGGCTGTGCGCAGAGAAACAACTGGTGTGACA 1628
Db      455  euIleGlyArpThrSerTrpGlySerGlyCysAlaLysValAlaTyrArgProGlyValTyr 475
      1629  CCAAGAGACAGAGTCTTCCCTGATTTACAGACAGATGAGAGAGAG 1678
Db      475  lvaAlaValMetValPheThrAspTrpIleTyrArgGlnMetAlaGlyAlaAsp 491

```

```

; EARLIER FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-342-749-2

Alignment Scores:
Pred. No.: 7.7e-50 Length: 492
Score: 768.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: Gaps: 15

US-10-806-370-11 (1-1748) x US-09-342-749-2 (1-492)

OY      288  GGTCCATCCGAGGAGTCAAT---CATCCGACAGTCAAGCTCGGTGACAACTCCCA 344
Db      12  GLyProTyrTyrGluValAsnHleGlyTyrGlnProGlnAsnPro-----TyrProAlaGln 29
OY      345  CCAGAGTGT-----ACCTTGTTAGACAAACACCAAGTGGAGGCTGTACCCATCC----- 392
Db      30  ProThrValValProThrValAlaTyrGlnValHleProAlaGlnTyrTyrProSerProVal 49
OY      393  GATATATCTCTCCGAGGTCAKACACCAACCAAGGCGCCAGGAGAGACCCAGGTACGA 452
Db      50  ProGlnTyrAlaProAlaGlyValLeuThrGlnAlaSerAsnPro-----ValValCys 66
OY      453  GCGTCCCAAGTTACCTGGCGGAGGCGCAAGAACAGTACCGCTCATCGGCTGCGTGC 512
Db      67  ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysValLeuCys 86
OY      513  -----TCCCTCATTTGCCCTGG 530
Db      87  lLeThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 106
OY      531  TGGTTTGCATCATCATCTCTTCCAGTGTGGCAGGCGCCACACAGGATCAGGTACAG 590
Db      107  LysPheMetGlySerLysCysSerAsnSerGlyIleGlyCysArpSerSerGlyThr--- 125
OY      591  AGCAGAGGAGAGAGCTGTCACAGACGCTGTGCTGAGCGGAGTGGATGACAGC 650
Db      126  -----CysIleAsnProSerAsnTrp--CysArpGlyValSerHleCysProG 141
OY      651  TGAAGATGACGAGCTGGCTGGCTGAGTGTGATCTGGACAAGTCTCTCTTAATAATCT 710
Db      141  lylGlylValArpGluValAsnArgCysValAlaArgLeuTyrGlyProAsnPheIleLeuGlnVal 161
OY      711  ACTGTGGCTCTCCCATCACTAGTGCCTTCCATCTGTGACACCAACTGATGACTCTACT 770
Db      161  ySerSerGlnArgLysSerTrpHleProValCysGlnArpArpTrpArgGlnValAsnTyrG 181
OY      771  CAGAGAACCTCCAGACACTGGAGTGGATTTGAGAGTGTCAACGAGAACAGGAGTTGCC 830
Db      181  lValGlnAlaAlaCysArgArpMetGlyTyrLysAsnAsnPheTyrSerSerGln--GlyI 200
OY      831  ACAGGATTTTGGCAACACTTCTCAATCTTGAGATCAATCCAC----- 877
Db      200  leValleArpSerSerGlySerThrSerPheMetLeuValAsnThrSerAlaGlyAlaValA 220
OY      878  --ATCCAGAAAGCTTCACAGGTCTGAA---TGCCCTTCACCGCGATATCTCCCTCC 932
Db      220  splleTyLysLeuValSerThrSerPheAlaCysSerSerLeuValValSerLeuA 240
OY      933  AGTGTTCACCTGCGAGCTGAGG-----GCCATGACCGCGCGATCTGGAGAGGG 983
Db      240  rgCysAlaCysGlyAlaLeuValLeuValSerSerArgGlnSerArgIleValGlyGly 260
OY      984  CGCTGCGCTCCGATAGCAAGTGGCTTGGCAAGTGAAGTGTGACTTCCGACCCACCA 1043

```

Db 260 luseralaleuProglYalatrProtrpInValserLeuHleValGlnasnValHisv 280
QY 1044 TCTGTGAGGACAGCCTATTGAGCCCGAGGGGTCTCATCGCCCGCACTGCTTCTTCG 1103
Db 280 a1CysglYgIyserIleIleThrProglutPrIleValThrAlaAlaHisCysVal---- 298
QY 1104 TGACCCGGGAGAGGCTCGTAGGGC-----TGGAGGTCTAGCGGGGACAGCA 1154
Db 299 -----GluYsProLeuasnAsnProtrpHisThrAlaPheAlaGlyIleLeuA 316
QY 1155 ACCTGACACAGTTCCTGAGGACCC-----TCCATTGCCGAGATCATCATCAAGCA 1208
Db 316 rglInserPheMetPheTyrGlyAlaGlyTyrGlnValGlnValIleSerHisProA 336
QY 1209 ATTACACCATGAGGAGAGCATATGACATCGCCCTTCATGCGGCTGTCCAGCCCTGA 1268
Db 336 anYrAspSerLysThrLysasnAsnPrIleAlaLeuMetLysLeuGlnLysProLeuT 356
QY 1269 CCCGTGCGGTCAATCCACCTGCTTGCCTCCCGCATGACATGACAGACCTTAGCCTCA 1328
Db 356 hrPheasnAsnLeuValLysProValCysLeuProasnProglYMetMetLeuGlnProG 376
QY 1329 ATGAGACCTGCTGATCAGAGCTTTGGCAAGACAGGAGACAGATGACAGATCC 1388
Db 376 luginLeuCystrPrIleSerGlyTyrPrIleAlaThrGlnGlu---LysGlyLysThrSerG 395
QY 1389 CCTTCTCCGGAGAGTGCAGGTCAATGTCATGCACTTCAGAAATGCAATGATCACTTGG 1448
Db 395 lValLeuasnAlaAlaLysValLeuLeuIleGlnThrGlnArgCysAsnSerArgTyrV 415
QY 1449 TCTATGACATTCATCTTACCCCAAGATGATGTGTGTGGGACCTTCGTGGGGGACAG 1508
Db 415 a1YrAsnAsnLeuIleThrProAlaMetIleCysAlaGlyPheLeuGlnGlyAsnValA 435
QY 1509 ACTCTCCAGGAGAGACAGCGGGGCTCTTGTCTGTGACAGAAACCGCTGTACC 1568
Db 435 spserCySgInglYAspSerGlyGlyProLeuValThrserLysAsnAsnIleTrpIle 455
QY 1569 TGGCAGGTGTACAGCTGGGGGACAGGCTGTGGCCAGAAACAACTGTGTGTACA 1628
Db 455 eulIleGlyAspThrSerTrpGlySerGlyCysAlaLysAlaTyrArgProglYalTyrG 475
QY 1629 CCAAGGACAGAGTCTTCCCTGGATTTACACCAAGTGAAGGAGAG 1678
Db 475 lYasnValMetValPheThrAspTrIleTyrArgGlnMetArgAlaAsp 491
RESULT 9
US-09-691-840-2
; Sequence 2, Application US/09691840
; Patent No. 644419
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TPRS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/691,840
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US/09/342,749
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/091,044
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-691-840-2
Alignment Scores: 7.7e-50 Length: 492

Score: 768.50 Matches: 178
Percent Similarity: 50.6% Conservative: 74
Best Local Similarity: 35.7% Mismatches: 195
Query Match: 23.1% Indels: 51
DB: 2 Gaps: 15
US-10-806-370-11 (1-1748) x US-09-691-840-2 (1-492)
QY 288 GGTTCATCTCCGAGCTAT---CATCCGACAGTCAAGCTCGGTACAACTTCCCA 344
Db 12 GlyProTyrTrpGlnAsnHisGlyTyrGlnProGlnAsnPro-----TyrProAlaGln 29
QY 345 CCAAGAGT-----ACCTGTAGAGCAACACAGTGGGGGTGACCATCC----- 392
Db 30 ProThrValValProThrValTyrGlnValHisProAlaGlnTyrTyrProserProVal 49
QY 393 GATCATCTCTGCGCAGTGCAGACCAAGCAACAGGGCCACAGGAGAGCCAGTACGA 452
Db 50 ProGlnTyrAlaProArgValLeuThrGlnAlaSerAsnPro-----ValValCys 66
QY 453 GCTGCCCAAGTTCACCTGGCGGGAGGCGCAAGACAGTACCGCTCATCGGCGCTGC 512
Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysValAlaLeuCys 86
QY 513 -----TCTCTCATTTGCCCTGG 530
Db 87 lIleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTrp 106
QY 531 TGGTTGCGCTCATCATCTCTCTCCAGTTCGTGAGGCGCACAGGAGATGATGATCAAG 590
Db 107 LysPheMetCylserLysCysSerAsnSerGlyIleGlnCysAspSerSerGlyThr--- 125
QY 591 AGCAGAGGAGACGTGCCCAAGACACCGCTTGCCTGTACCGGGGTGTGACTGCAGC 650
Db 126 -----CysIleAsnProSerAsnTrp---CysAspGlyValSerHisSerProG 141
QY 651 TGAAGAGTGCAGAGCTGGCGGTGAGGTTGACTGTGAGCAAGTCTGTAAATCT 710
Db 141 lYelGlnAspGlnAsnAsnGlyValAlaGlyLeuTyrGlyProAsnPheIleLeuGlnValT 161
QY 711 ACTCTGGGTCCTCCATCATGAGTGTCCCATCTGTAGACAACTGATGATCTTACT 770
Db 161 yIserSerGlnAlaGlySerTrpHisProValCysGlnAspAspTrpAsnGlnAsnTyrG 181
QY 771 CAGAGAACCTGCAGACAGCTGGGTTGAGAGTGTCTCACCGGACACAGGAGTGGCC 830
Db 181 lYArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
QY 831 ACAGGATTTGGCAACGCTTCTCAATCTGAGATCACTCCAC----- 877
Db 200 lValAspAspSerGlySerThrPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
QY 878 --ATCCAGAAAGCTCCACAGGTGAA---TGCCTTCCACAGCGGTATATCTCCCTCC 932
Db 220 spIleTyrLysLysLysLeuTyrHisSerThrAspAlaCysSerSerLysAlaValValSerLeuA 240
QY 933 AGTGTTCACCTGGGAGTGAAG-----GCCATGACCGGGCGGATGTGGAGGGG 983
Db 240 rGcysIleAlaCysGlyValAlaAsnLeuAsnSerSerArgInserArgIleValGlyG 260
QY 984 CGCTGGCTGGATGACAGAGTGTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1043
Db 260 luseralaleuProglYalatrProtrpInValserLeuHleValGlnasnValHisv 280
QY 1044 TCTGTGAGGACAGCCTATTGAGCCCGAGGGGTCTCATCGCCCGCACTGCTTCTTCG 1103
Db 280 a1CysglYgIyserIleIleThrProglutPrIleValThrAlaAlaHisCysVal---- 298
QY 1104 TGACCCGGGAGAGGCTCTGAGGGC-----TGGAGGTCTAGCGGGGACAGCA 1154
Db 299 -----GluYsProLeuasnAsnProtrpHisThrAlaPheAlaGlyIleLeuA 316
QY 1155 ACCTGACACAGTTCCTGAGGACCC-----TCCATTGCCGAGATCATCATCAAGCA 1208


```
Db      376 lucinleucystrpilesersglytrpdlathngluh---lyeglyyethrsersg 395
      1389 CCTTCCTCCGGAGGTGAGGTGATCTCATCGACCTTCAGAAATGCAATGACTACTTGG 1448
      395 luvalleuamnalalyleuvalleuuleuilegltuthnglnargcyasnsersargtyrv 415
      1449 TCTATGACAGTTACTTACCCCAAGATGATGTGTGTGGAGACTTTCGTGGGGCAGAG 1508
      415 altyrbaenleuilethtrproalmetleicybalaglyphleuendlnlyamvala 435
      1509 ACTCTGCGCAGGAGACAGCGGGGGCTCTTGTCTGTGACAGAAACAACCGCTGGTACC 1568
      435 spsercygngllyaspsersglytylproleuvalthnserasnsamnsenlettrptpl 455
      1569 TGGCAGAGGTGACAGCTGGGGCAGACAGCTGTGGCAGAGAAACAACCTGGTGTACA 1628
      455 euileglyasrpthserrtpolysersglycybalalyalaatyrtargproglyvaltyrg 475
      1629 CCAAGTGACAGAAATTCTTCCTGATTTACAGCAAGATGAGAGC 1675
      475 lyaenvalmetvalphenethrasprtrplettyrtarglnmetlyala 490
```

```
RESULT 12
US-09-879-792-14
; Sequence 14, Application US/09879792
; Patent No. 6734006
; GENERAL INFORMATION:
; APPLICANT: Xliao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-879-792-14 (Docket No. 6734006 LIO-81-MO)
```

```
Alignment Scores:
Pred. No.: 1,55e-49 Length: 492
Score: 764.50 Matches: 178
Percent Similarity: 50.5% Conservative: 73
Best Local Similarity: 35.8% Mismatches: 195
Query Match: 22.9% Indels: 51
Gaps: 15
```

US-10-806-370-11 (1-1748) x US-09-879-792-14 (1-492)

```
QY      288 GGTCTTCATCCGAGGATCAT---CATCCGCGAGGTGAGCTCGGTGACAACTCCCA 344
      12 glytrprotyrtyrgluamhileglytyrqlnprogluampro-----tyrproalagln 29
      345 CCAAGAGT-----ACCTGTATGAGCAACAACAGTGGGGCGGTACCACTCC----- 392
      30 ProthrValValProthrValTyrglyValHisProalaglnTytyrProserProval 49
      393 GATCATCTCCGCGAGGTGACAGACCAAGCAAGGAGCCCAAGGAGAGGCCAGGTACGA 452
      50 ProglntyrAlaProAlaTyrglyValleuThrGlnAlaSerAmPro-----ValValCys 66
```

```
QY      453 GCCTCCCAAGTTCACTGGGAGGAGGCGCAAGACACTACCGCTCATCGGGTGGCTGC 512
      67 ThrclnProlysserserProserglyThrValCysethrserythnlyblyalaleucys 86
      513 -----TCCTCTCATTCGCTGG 530
      87 lIethrleuthrleuglythrphenleuvalglyalalaleuvalalaglyleuvaltrp 106
      531 TGGTTGCTCATCATCTCTTCAGTTCTGGCAGGGCCACACAGGATGAGTACAGAG 590
      107 lyphenmetglyserlybCyssersamserclylleuCyaspsersersglythr--- 125
      591 AGCAGAGGAGAGGTGTCCCAAGCAGCTGTGGCTGTGACAGGGGTGTGACTGCAAG 650
      126 -----Cyelleamproserasnthr--CyasrglyValserthlecyserProg 141
      651 TGAAGACTGACAGCTGGCGCTGGTGTGAGTTTGACTGGACAAAGTCTGTGTAATCT 710
      141 lyglyluasrpgluasnarCyasvalargleutyrglylProamphenleleuendlnmet 161
      711 ACTCTGGGTCTCCATCATGAGCTTCCATCTGTATGACAGCAACTGAGTACTCTTACT 770
      161 ytserserglnarglybyserrtrphsiprovalCysglnasrparprtrpaendlnamtyrg 181
      771 CAGAGAACTTCGACAGCAGCTGGTTTGAGAGTGTCAACCGGACAAACGAGGTGCC 830
      181 lyargalalalacyargasrmercltyrtyrvalasnaenphenetyrserserdln---glyI 200
      831 ACAGAGATTTTGGCAACAGCTTTCATCTTGAGATACAACTCCACC----- 877
      200 levalasrparsegslyserthrsphenmetlybvalamnthrservalglyamvala 220
      878 --ATCCAGGAAGGCTCCACAGGTCTGAA--TGCCCTTCCAGCGGTATATCTCCCTCC 932
      220 eprietyrlyblybelythnhsersarphalacysserserlybalaValaserleua 240
      933 AGTGTTCACACTGGAGTACAG-----GCCATGACCGGGCGGATGTGTGAGAGGG 983
      240 rgyblyeualacybglyvalasneleuansersersarglnserargllevalglylg 260
      984 CGCTGGCTCGGATAGCAAGTGTCTGGAGAGTGTGCACTTGCGGACCCAGCA 1043
      260 luserlalaleuproglalatrptrottrpoinvalserleuhsvalaglnamvalahsv 280
      1044 TCTGTGAGGACAGCTCATTTGACGCGGAGTGTGTCTCATGCGCCCACTGTCTTCTGG 1103
      280 alCybglylyserllethtrproglutrpilvalthralaleahleicyvala1---- 298
      1104 TGACCCGAGAGAGTCTTGAGGGC-----TGAAAGTGTACGGGAGCCAGCA 1154
      299 -----gluLyProleuamnsanProtrphsistrphtalalephalaglylleuua 316
      1155 ACCTGACACAGTTCCCTGAGGAGCC-----TCCATTCGCGAGATGATATCAACAGA 1208
      316 rglinsersphenemphenetyrglyalaglytyrglnvalaglnlyvalalbsertlssproa 336
      1209 ATTACACGATGAGAGAGAGACTATGACATCGCGCTCATGCGGCTGTGCAAGCCCTGGA 1268
      336 entyrtasrserlybthlyasamnsanhsprllealaleumetylvalendlnlyseroleut 356
      1269 CCTGTCCGCTCACATCCACCTGCTTGTCTCCATGATGATGACAGACACTTTAGCTCA 1328
      356 hrphenamleuvaltyrrvalalysleuproamprogllymetmetleuendlnprog 376
      1329 ATGACAGCTCTGTGATACAGCTTTGGCAAGACCAAGGAGACAGATGACAAAGCATCC 1388
      376 lucinleucystrpilesersglytrpdlathngluh---lyeglylyethrsersg 395
      1389 CCTTCCTCCGGAGGTGAGGTGATCTCATCGACCTTCAGAAATGCAATGACTACTTGG 1448
      395 luvalleuamnalalyleuvalleuuleuilegltuthnglnargcyasnsersargtyrv 415
      1449 TCTATGACAGTTACTTACCCCAAGATGATGTGTGTGGAGACTTTCGTGGGGCAGAG 1508
```


Db 415 a1TyrAspAsnLeu1leThrProAlaMet1leCyAlaGlyPheLeuGlnGlyAsnValA 435
QY 1509 ACTCTGTCAGAGAGAGACAGCGGGGCTTGTCTGTGAGACAAACCGCTGTACC 1568
Db 435 epsrCyvGlnGlyAspSerGlyGlyProLeuVal1ThrSerAsnAsn1leTrrpL 455
QY 1569 TGGCAGGTGTACCAAGCTGGGGGACAGGCTGTGGCCAGAAACAACTGTGTGTA 1628
Db 455 eul1eGlyAspThrSerTrpGlySerGlyCyAlaValAlaTyrArgProGlyValTyrG 475
QY 1629 CCAAGCAGACGAAGTCTTCCCTGGATTTCACCAAGATGGAAGC 1675
Db 475 lyaAsnValMetValPheThrAspTrpIleTyrArgGlnMetLysAla 490

RESULT 14
US-09-759-143-895
Sequence 895, Application US/09759143
Patent No. 6800746
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: PaetSEQ for Windows Version 3.0
SEQ ID NO 895
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-895

Alignment Scores:
Pred. No.: 1,55e-49 Length: 492
Score: 764.50 Matches: 178
Percent Similarity: 50.5% Conservative: 73
Best Local Similarity: 35.8% Mismatches: 195
Query Match: 22.9% Indels: 51
Gaps: 15

US-10-806-370-11 (1-1748) x US-09-759-143-895 (1-492)

QY 288 GGTCTCATCGGAGAGTCA---CATCCGCGAGGTACAGCTGGTGAACACTCCCAA 344
Db 12 G1yProTyrTyrGluAsnH1sG1yTyrGlnProGlnAsnPro-----TyrProAlaGln 29
QY 345 CCAGAGTGT-----ACCTTGTTAGAGCAACACAGTGGGGGCTGTACCAATCC----- 392
Db 30 ProThrValValPProThrValTyrGlnValH1sProAlaGlnTyrTyrProSerProVal 49
QY 393 GATCATCTCTCCAGGTCAGACACCAACCAAGGCGCCACCAAGGAGAGCCCAAGTACGA 452
Db 50 ProGlnTyrAlaProAlaProAlaValLeuThrGlnAlaSerAsnPro-----ValValCys 66
QY 453 GCGTGGCCAAAGTTTACCTGGCGGGAGGGGACCAAGAACGCTTACCGCTCATGGGCTGCG 512

Db 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysLysAlaLeuCys 86
QY 513 -----TCCTCTCATATGCGCTGG 530
Db 87 l1eThrLeuThrLeuGlyThrPheLeuValG1yAlaAlaLeuAlaAlaGlyLeuLeuTrrp 106
QY 531 TGGTTTGCTCATCATCTCTTCCAGTCTGGCAGGGCCACACAGGATCAGGTACAG 590
Db 107 LysPheMetGlySerLysCysSerAsnSerGly1leGlnCysAspSerSerGlyThr--- 125
QY 591 AGCAGAGAGAGAGCTGCCAAGACACCTGTCCCTGTGACGGGGTGTGACTGCACAG 650
Db 126 -----Cys1leAsnProSerAsnTrp---CysAspGlyValSerH1eCysArg 141
QY 651 TGAAGATGACAGAGCTGGGCTGCGTGGAGTTTGACTGGACCAAGTCTGTCTTAAATCT 710
Db 141 lylGlyGluAspGlnAsnArgCysValAlrGlyLeuTyrGlyProAsnPhe1leLeuGlnMet 161
QY 711 ACTGTGGGTCTCCCATCAAGTGGCTTCCCATCTGTAGCAGACCAATGGAATGACTCTACT 770
Db 161 YrSerSerGlnArgLysSerTrpH1sProValCysGlnAspAspTrpArgGlnAsnTyrG 181
QY 771 CAGAGAGACCTGCGCAGACAGCTGGGTTCCAGAGTCTCACCGGACCAACGAGTTGCC 830
Db 181 lyaAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
QY 831 ACAGGATTTTGGCAACAGCTTCCATCTTGAGATACAACTCCACC----- 877
Db 200 leValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerAlaGlyAsnValA 220
QY 878 --ATCCAGAAAGCTCCACAGGTTGAA--TGCCCTTCCACCGGTATATCTCCCTCC 932
Db 220 sp1eTyrLysLysLeuTyrH1sSerAspAlaCysSerSerLysAlaValAlaSerLeuA 240
QY 933 AGTGTCCCACTGGGACTAGC-----GCCAGACCGGGCGGATGTGGAGAGG 983
Db 240 rGysLeuAlaCysGlyValAlaAsnLeuAsnSerSerArgGlnSerArg1yGly 260
QY 984 CGCTGGCTCGGATAGCAATGGGCTTGGCAGAGTGAAGTGTGCACTTGGCAGCCACCA 1043
Db 260 luseAlaLeuProGlyAlaTrrpTrpTrpAlaValSerLeuH1sValGlnAsnValH1sV 280
QY 1044 TCTGTGAGGAGCGCTCATTTGACGCCAGTGGTGTCTACTGCGGCCACTGCTTCTCG 1103
Db 280 alCysGlyGlySer1le1leThrProGlnTrp1leValTrnAlaAlaH1eCysVal---- 298
QY 1104 TGACCCGGAGAAAGTCTTGAGAGCC-----TGAAGGTGTAGCGGGACCAAGCA 1154
Db 299 -----GluLysProLeuAsnAsnProTrrpH1sTrrpAlaPheAlaGly1leLeuA 316
QY 1155 ACCTGCACACAGTTGCTGAGAGCAGCC-----TCCATTTGCCAGATCATCAACAGCA 1208
Db 316 rGlnSerPheMetPheTyrGlyAlaGlyTyrGlnValGlnLysVal1lSerH1sProA 336
QY 1209 ATTACACCGATGAGAGAGACGACTATGACATCGCCCTCATGCGCTGTCCAAAGCCCTGA 1268
Db 336 snTyrAspSerLysThrLysAsnAsnAsp1leAlaLeuMetLysLeuGlnLysProLeuT 356
QY 1269 CCTGTCCGCTCATCATCAACCTGCTGCTCCCTCCCATGATGACAGACACCTTTAGCTCA 1328
Db 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
QY 1329 ATGAGACCTCTGGATCAACAGGCTTGGCAAGACCAAGAGAGACAGATGACAAAGCATCCC 1388
Db 376 luginLeuCyStrp1leSerGlyTrrpGlyAlaThrGlnGlu---LysGlyLysThrSerG 395
QY 1389 CCTTCTCCGAGAGGTGACAGTCAATCTCATGCACTTGCAAGAAATGCAATGACTACTGG 1448
Db 395 luvAlleuAsnAlaAlaLysValLeuLeuTrrpGlnArgCysAsnSerArgTrrv 415
QY 1449 TCTATGACAGTTCATCTTACCCCAAGATGATGTGTGTGGGACCTTTCGTGGGGGAGAG 1508
Db 415 alTyrAspAsnLeu1leThrProAlaMet1leCyAlaGlyPheLeuGlnGlyAsnValA 435

QY 1509 ACTCTGCGCAGGAGACAGAGGGGGCTCTGCTGTGAGCAACACCGCTGTACC 1568
 DB 435 spherCysGlnGlyAspSerGlyGlyProLeuValThrSerAsnAsnAsnIleTyrIle 455
 QY 1569 TGGCAGGTGTCCACGAGCTGGGGGACAGAGCTGTGGCCAGAAACAACTGTGTGTACA 1628
 DB 455 euleGlyAspThrSerTyrGlySerGlyCysAlaLysAlaTyrArgProGlyValTyrG 475
 QY 1629 CCAAGTGACAGAAAGTTCTTCCCTGATTTACAGCAAGATGAGAGC 1675
 DB 475 LysValMetValPheThrAspTyrIleTyrArgGlnMetLysAla 490
 RESULT 15
 US-10-012-896-895
 / Sequence 895, Application US/10012896
 / Patent No. 6943236
 / GENERAL INFORMATION:
 / APPLICANT: Xu, Jiangchun
 / APPLICANT: Dillon, Devin C.
 / APPLICANT: Mitcham, Jennifer L.
 / APPLICANT: Harlocker, Susan L.
 / APPLICANT: Jiang, Yugu
 / APPLICANT: Kalos, Michael D.
 / APPLICANT: Retter, Marc W.
 / APPLICANT: Stolk, John A.
 / APPLICANT: Day, Craig H.
 / APPLICANT: Vedvick, Thomas S.
 / APPLICANT: Carter, Darick
 / APPLICANT: Li, Samuel X.
 / APPLICANT: Wang, Aijun
 / APPLICANT: Skeiky, Yasar A.W.
 / APPLICANT: Hepler, William T.
 / APPLICANT: Henderson, Robert A.
 / APPLICANT: Hurst, John
 / APPLICANT: McNeill, Patricia D.
 / APPLICANT: Houghton, Raymond L.
 / APPLICANT: Vinals de Basols, Carlota
 / APPLICANT: Foy, Teresa
 / APPLICANT: Fanger, Gary R.
 / APPLICANT: Mantanabe, Yoshihiro
 / APPLICANT: Mesgher, Madeleine Joy
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 / FILE REFERENCE: 210121.427C27
 / CURRENT APPLICATION NUMBER: US/10/012.896
 / CURRENT FILING DATE: 2001-12-10
 / NUMBER OF SEQ ID NOS: 1011
 / SOFTWARE: FASTSEQ for Windows Version 3.0
 / SEQ ID NO 895
 / LENGTH: 492
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-10-012-896-895
 Alignment Scores:
 Pred. No.: 1.55e-49 Length: 492
 Score: 764.50 Matches: 178
 Percent Similarity: 50.5% Conservative: 73
 Best Local Similarity: 35.8% Mismatches: 195
 Query Match: 22.9% Indels: 51
 DB: 2 Gaps: 15
 US-10-806-370-11 (1-1748) x US-10-012-896-895 (1-492)
 QY 288 GGTCTTATCCGGGAGGTGAT---CATCCGCGAGGTACGCTCGGTGACACCTCCCA 344
 DB 12 GtlyProtyrTyrGluAsnHisGlyTyrGlnProGluAsnPro-----TyrProAlaGln 29
 QY 345 CCAGAGTGT-----ACCTGTAGACCAACAGAGTGGGGCTGTACCATCC----- 392
 DB 30 ProthrValValProThrValTyrGlnValHisProAlaGlnTyrTyrProSerProVal 49

QY 393 GATCATCTCTCCAGAGTACGACCAACAGAGCCACAGAGAGAGCCAGGTACGA 452
 DB 50 ProGlnTyrAlaProAlaGlyValLeuThrGlnAlaSerAsnPro-----ValAlaCys 66
 QY 453 GCCTGCCCAAGTTCACTTGGCGGGAGGCGCAAGACGCTACCGCTCATCGGGTGTGC 512
 DB 67 ThrGlnProLysSerProSerGlyThrValCysThrSerLysThrLysAlaLeuCys 86
 QY 513 -----TCCTCCATTCCTGCTG 530
 DB 87 IleThrLeuThrLeuGlyThrPheLeuValGlyAlaAlaLeuAlaAlaGlyLeuLeuTyr 106
 QY 531 TGGTTTCGCTCATCATCTCTTCCAGTTCTGCGAGGCGCCACAGAGGATCAGTACAG 590
 DB 107 LysPheMetGlySerLysCysSerAsnSerGlyIleGluCysAspSerSerGlyThr--- 125
 QY 591 AGCAGAGGAGAGAGCTGTCCCAACAGCAGCTGTGCTGTGACGGGGTGTGACTGCAC 650
 DB 126 -----CysIleAsnProSerAsnTyr--CysAspGlyValSerHisCysProG 141
 QY 651 TGAAGAGTACAGAGCTGGGCTGGGTGAGTTGACTGGACAGAGTCTCTGCTTAATCT 710
 DB 141 LysGlyGluAspGluAsnArgCysValArgLeuTyrGlyProAsnPheIleLeuGlnMet 161
 QY 711 ACTGTGGTCTCCCATCATGAGCTTCCATCTGTAGCAGCAACTGGAATGACTCTACT 770
 DB 161 TyrSerSerGlnArgLysSerTyrPheIleProValCysGlnAspAspTyrAsnGlnAsnTyrG 181
 QY 771 CAGAGAAAGCTGCGCAGACAGCTGGTTTCGAGAGTGTCAACCGGACAGAGGTGCC 830
 DB 181 LysArgAlaAlaCysArgAspMetGlyTyrLysAsnAsnPheTyrSerSerGln---GlyI 200
 QY 831 ACAGGATTTTGGCAACAGCTTCTCATCTTGAGATCAACTCCAC----- 877
 DB 200 LeValAspAspSerGlySerThrSerPheMetLysLeuAsnThrSerValGlyAsnValA 220
 QY 878 --ATCCAGAAAGCTCCACAGGCTGTA---TGCCCTTCCAGCGGTATATCCCTCC 932
 DB 220 splleTyrLysLysLeuTyrHisSerAspAlaCysSerSerLysAlaValValSerLeuA 240
 QY 933 AGTGTTCACATCGGAGCTGAGG-----GCCATGACCGGGCGGATCTGTGGAGGG 983
 DB 240 rGysLeuAlaCysGlyValAsnLeuAsnSerSerArgGlnSerArgIleValGlyIlg 260
 QY 984 CGCTGGCTCGGATAGCAAGATGCGCTTGGCAAGTGAATCTGCACCTTGGCACCCACA 1043
 DB 260 InsSerAlaLeuProGlyAlaTyrProTyrGlnValSerLeuHisValGlnAsnValHisv 280
 QY 1044 TCTGTGAGGACAGCTGATTGAGCGCCAGTGGGCTCCATCGCGCCACTGCTTCTTG 1103
 DB 280 alCysGlyLysSerIleIleThrProGlnTyrIleValThrAlaAlaHisCysVal---- 298
 QY 1104 TGACCCGGAGAAAGTCTGAGAGGC-----TGAAGGTGTACCGGGCACCCAGA 1154
 DB 299 -----GluLysProLeuAsnAsnAsnProThrPheIleThrAlaPheAlaGlyIleLeuA 316
 QY 1155 ACCTGACCAAGTGTCTGAGCAGCC-----TCCATGTCCGAGATCATCAACAGCA 1208
 DB 316 rGlnInsPheMetPheTyrGlyAlaGlyTyrGlnValGlnLysValIleSerHisProA 336
 QY 1209 ATTTACACCATGAGAGGAGAGCAATGACATCGCGCTCATGCGGCTGTGCCAAGCCCTGA 1268
 DB 336 snTyrAspSerLysThrLysAsnAsnAspIleAlaLeuMetLysLeuGlnLysProLeuT 356
 QY 1269 CCTGTCCGCTCATCAACCTCTGCTTCCCTCCCATGATGACAGACCTTTAGCCTCA 1328
 DB 356 hrPheAsnAspLeuValLysProValCysLeuProAsnProGlyMetMetLeuGlnProG 376
 QY 1329 ATGAGACCTGTGTGATACAGGCTTTGGCAAGACAGGAGAGACAGATGACAAAGCATCC 1388
 DB 376 IugInLeuCystrIleSerGlyTyrPheIleAlaThrGlnGlu---LysGlyLysThrSerG 395
 QY 1389 CCTTCTCCGGAGAGGTACAGGTCAATCTCATGCACTTCAAGAAATGCAATGACTTGG 1448

Db 395 luValleuasnAlaAlaIaIaValleuenuIeGIuThnGlnArgCYabaSbnSerArgItyrV 415

QY 1449 TCTTATGACAGCTTACCTACCCCAAGAGATAGTGTGTGGGGAACCTTGTGTGGGCGACAG 1508

Db 415 aIlyrAspAsnneuIeIthrProAlaMetIeCYabaGIyPheueIdmIlysnValA 435

QY 1509 ACTCTGTGCACAGGAGACAGCGGGGGCCCTTGTGTGTGACACAGAACCGCTGTGACC 1568

Db 435 spSerCYsGInGIyAspserGIyGIyProIeuValIthrSerAsnAsnIleItrpIPL 455

QY 1569 TGGAGAGGTATCACAGCTGTGGGCGACAGGCTGTGGCCAGAGAAACAACTGTGTGTACA 1628

Db 455 euIleGIyAspThrSerItrpGIyserGIyCYabaIyAsaIyArgProIlyValIyArg 475

QY 1629 CCAAGTGCACAGAAAGTCTTCCCTGGAGATTTACACAGATGTGAAGC 1675

Db 475 IyAsnValMetValPheThrAspItrIeIyArgGInMetIyValA 490

Search completed: September 16, 2006, 01:56:47
Job time : 93 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 16, 2006, 02:28:48 | Search time 58.1 Seconds
(without alignments)
4180.893 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 ctcagagaccatcgagagaggg.....ggctgtgtgactcgagagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2097797 seqs, 463214858 residues
Total number of hits satisfying chosen parameters: 4195594

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1p
-O=/abs/ABSSWB.spool/US10806370/runat.15092006.105646.10654/app_query.fasta_1
-DB=Published Applications AA Main -OEMT=fastan -SUFFIX=n2p.rapbm
-MIMATCH=0.1 -LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=ptco -NORM=ext
-HEAPSIZE=500 -MNLN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10806370 @CCN 1.1 446 @runat.15092006.105646.10654 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA Main:
1: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US10_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/pcodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3004	90.1	569	4	US-10-428-275-130
2	2999	90.0	562	3	US-09-879-792-12
3	2999	90.0	562	4	US-10-806-370-12
4	2991	89.7	562	4	US-10-156-214A-24
5	2991	89.7	562	5	US-10-916-758-94
6	2963	88.9	688	4	US-10-156-214A-26
7	2963	88.9	688	5	US-10-916-758-2
8	2956	88.7	581	4	US-10-353-690-100
9	2953.5	88.5	586	4	US-10-428-275-120
10	2948.5	88.5	586	4	US-10-428-275-104
11	2948.5	88.5	586	4	US-10-428-275-124

12	2943.5	88.3	586	4	US-10-428-275-142	Sequence 142, App
13	2942.5	88.3	586	4	US-10-428-275-140	Sequence 140, App
14	2775.5	83.3	537	3	US-09-888-615-104	Sequence 104, App
15	2775.5	83.3	537	4	US-10-428-275-128	Sequence 128, App
16	2775.5	83.3	537	6	US-11-037-243-104	Sequence 104, App
17	2691	80.7	542	4	US-10-428-275-122	Sequence 122, App
18	2291.5	68.8	477	4	US-10-177-661-2	Sequence 2, App1
19	2291.5	68.8	477	5	US-10-910-507-2	Sequence 2, App1
20	2287	68.6	421	3	US-09-898-837A-15	Sequence 15, App1
21	2287	68.6	421	4	US-10-428-275-132	Sequence 132, App
22	2219.5	66.6	466	4	US-10-274-639-8	Sequence 8, App1
23	2219.5	66.6	466	4	US-10-333-574-8	Sequence 8, App1
24	2130.5	63.9	446	4	US-10-428-275-144	Sequence 144, App
25	2108	63.2	446	4	US-10-177-661-4	Sequence 4, App1
26	2108	63.2	446	5	US-10-910-507-4	Sequence 4, App1
27	2097	62.9	382	4	US-10-428-275-134	Sequence 134, App
28	2059	61.8	406	4	US-10-428-275-126	Sequence 126, App
29	2059	61.8	406	4	US-10-428-275-138	Sequence 138, App
30	2026	60.8	401	4	US-10-428-275-112	Sequence 112, App
31	2016	60.5	401	4	US-10-428-275-116	Sequence 116, App
32	2012	60.4	401	4	US-10-428-275-118	Sequence 118, App
33	1842.5	55.3	370	4	US-10-428-275-114	Sequence 114, App
34	1839.5	55.2	367	4	US-10-428-275-110	Sequence 110, App
35	1474	44.2	283	4	US-10-428-275-148	Sequence 148, App
36	1311	39.3	296	3	US-09-804-156-30	Sequence 30, App1
37	1311	39.3	296	4	US-10-067-761-30	Sequence 30, App1
38	1311	39.3	296	4	US-10-319-519-30	Sequence 30, App1
39	1311	39.3	372	3	US-09-804-156-15	Sequence 15, App1
40	1311	39.3	372	3	US-09-946-633-7	Sequence 7, App1
41	1311	39.3	372	4	US-10-125-459-7	Sequence 7, App1
42	1311	39.3	372	4	US-10-067-761-15	Sequence 15, App1
43	1311	39.3	372	4	US-10-319-519-15	Sequence 15, App1
44	1270	38.1	239	4	US-10-428-275-106	Sequence 106, App
45	1270	38.1	239	4	US-10-428-275-108	Sequence 108, App

ALIGNMENTS

RESULT 1
US-10-428-275-130
Sequence 130, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OR INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428, 275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseq1st version 0.1
SEQ ID NO 130
TYPE: PRT
ORGANISM: Homo sapiens
US-10-428-275-130
Alignment Scores:

Pred. No.: 8.36-175 Length: 569
Score: 3004.00 Matches: 563
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.1% Gaps: 0
DB: 4

US-10-806-370-11 (1-1748) x US-10-428-275-130 (1-569)

```
QY 8 ACCATGAGAGAGACAGCCAGGGAATGATCTCCAGCAAGAACCTTCAGTGAACA 67
DB 4 ThiretGluATGAsperSerHisGlyAsnAlaSerProAlaArgThrProSerHisGlyAla 23
QY 68 TCTCCAGCCAGGAGATCTCCAGCTGGAGACCTCCAGGCCGGGAGATCTCCAGCCAGGCA 127
DB 24 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 43
QY 128 TCTCCAGCCAGGAGATCTCCAGCTGGAGACCTCCGGGCCGGGAGATCTCCAGCCAGGCA 187
DB 44 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 63
QY 188 TCTCCAGCTGGTACCTCCAGGCCGGGAGATCTCCAGGCCGGGAGATCTCCAGCCAGGCA 247
DB 64 SerProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAla 83
QY 248 TCTCCAGCCAGGAGATCTCCAGCTGGAGACCTTCAGATCTTCAGGCTCTCATCCGAGAGTCA 307
DB 84 SerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerSerGlyArgSer 103
QY 308 TCATCCGCGAGAGTCAAGCTCGGTGACAACTTCCCAACAGAGTACCTTTGTTAGACA 367
DB 104 SerSerHisArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAla 123
QY 368 ACACGAGTGGGGGGGTACCCATCCGATCTCTGCGAGTGAAGACAGCAACAGG 427
DB 124 ThrProAlaGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArg 143
QY 428 GCCACAGGAGAGACCCAGGTACGAGCTGCGCAAGTTACCTGAGCGGAGAGGCGCAAG 487
DB 144 AlaThrArgGlnSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlyGlnLys 163
QY 488 CAGCTACCGCTCAATCGGGTGGTGGTCTCTCTCATTTGCCCTGGTGGTCTCTCATATC 547
DB 164 GlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValIleSerLeuIleIle 183
QY 548 CTCTTCCAGTCTGGCAGGAGCACAGAGATCAAGGTACAGGAGAGGAGAGAGCTG 607
DB 184 LeuPheGlnPheThrGlnGlyHisThrGlyIleArgGlyLysGlnArgGlnSerCys 203
QY 608 CCCAAGCAGCTGTTGCTGTGACCGGGTGGTGAATGCAAGCTGAAGAGTGAAGAGCTG 667
DB 204 ProLysHisAlaAlaValArgCysAspIlyValIleAspCysLysLeuLysSerAspGlnLeu 223
QY 668 GCGTGGCTGAGGTTGACTGGGACAAGTCTGCTTAAATCTACTCTGGTCTCTCCAT 727
DB 224 GlyCysValArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHis 243
QY 728 CAGTGGCTTCCATCTGTAGAGAGCAACTGGAATGACTCTTACTGAGAGAAACCTGGCAG 787
DB 244 GlnTrpLeuProIleCysSerSerSerAsnTrpAsnAspSerTyrSerGlnLysThrCysGln 263
QY 788 CAGCTGGATTTCAGAGTGTACACCGGCAACCGAGTTGGCCACAGGAGATTTTGGCAAC 847
DB 264 GlnLeuGlyPheGlnSerAlaHisArgThrArgValAlaHisArgAspPheAlaAsn 283
QY 848 AGCTTTCATCTTGAGATACAACTCCACATCCAGGAAAGCTTCACAGGCTTGAATGC 907
DB 284 SerPheSerIleLeuArgTyrAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCys 303
QY 908 CCTTCCGAGGGGTATATCTCCCTCCAGGTTCCCACTGGCGGAGCTGAAGGCGCATACCGGG 967
DB 304 ProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrIly 323
```

```
QY 968 CGAGTCTGGAGAGGCGCTGGGCTCGGATAGCAAGTGGCTTGGCAAGTGAATCTGCAC 1027
DB 324 ArgIleValIleGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValIleSerLeuHis 343
QY 1028 TTGGGACCCACCAATCTGTGAGAGGACGCTTATGACCCCAAGTGGGGTCTCACTGCC 1087
DB 344 PheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValIleThrAla 363
QY 1088 GCCCACTGCTTCTTCGTGACCCGGGAGAAAGTCTGAGGAGCTGGAAGGTGATACGCGGC 1147
DB 364 AlaHisCysPhePheValThrArgGlnLysValLeuIleGlnTrpLysValTyrAlaGly 383
QY 1148 ACCAGCAACTGACCAAGTTGCTGAGGACAGCTTCCATTCGCGAGATCATCAACAGC 1207
DB 384 ThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleLeuAsnSer 403
QY 1208 AATTACCGATAGAGAGAGCACTATGATATGCGCTCATGGGCTGTCCAGCGCTG 1267
DB 404 AsnTyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetCysLeuSerLysProLeu 423
QY 1268 ACCGTGCTCGCTCATCAACCTGCTTGGCTTCCCATGATGACAGACCTTTAGCTC 1327
DB 424 ThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeu 443
QY 1328 AATGAGACTGCTGGATCAAGGCTTTGGCAAGCCAGGAGACAGATGACAAACATCC 1387
DB 444 AsnGlnThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspLysThrSer 463
QY 1388 CCTTCTCGGGAGGAGTGCAGGTCAATCTCATGACTTCAAGAAATGCAATGACTTCTG 1447
DB 464 ProPheLeuArgGlnValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeu 483
QY 1448 GTCTATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGAGACTTCTGGGGGAG 1507
DB 484 ValTyrAspSerTyrIleLeuThrProArgMetCysAlaGlyAspLeuArgGlyArg 503
QY 1508 GACTCTGCGCAGGAGACAGCGGGGGCTCTTGTCTGTGACGCAACCAACCGTGGTAC 1567
DB 504 AspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyr 523
QY 1568 CTGGCAGTGTGACACGCTGGGAGCAGGCTGGGCGCAGAGAAACCAACTGGTGTGAC 1627
DB 524 LeuAlaGlyValThrSerTrpGlyThrGlyCysGlyAlaTrpAsnLysProGlyValTyr 543
QY 1628 ACCAAGTGAAGAAAGTCTTCCCTGGATTTTACAGCAAGATGAGAGAGAGTGCATTC 1687
DB 544 ThrIleValThrGlnValLeuProTrpIleTyrSerLysMetGlnSerGlnValArgPhe 563
QY 1688 AGAAATCC 1696
DB 564 ArgLysSer 566
```

RESULT 2
US-09-879-792-12
Sequence 12, Application US/09879792
Patent No. US20020061850A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
APPLICANT: Gedrich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Sertine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/09/879, 792
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211, 224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283, 353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283, 648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36

(Docket No. US20020061850A1 LIO-81-WO)

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 12

; LENGTH: 562

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-879-792-12

Alignment Scores:

Pred. No.:	1,67e-174	Length:	562
Score:	2999.00	Matches:	562
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	90.0%	Indels:	0
DB:	3	Gaps:	0

US-10-806-370-11 (1-1748) x US-09-879-792-12 (1-562)

```
QY 11 ATGAGAGGAGGAGCCAGGGAATGATCTCCAGCAAGAACACCTTCAGCTGGAGCATCT 70
    |||
Db 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
    |||
QY 71 CCAGCCAGGAGCATCTCCAGCTGGAGACATCCAGCCGGGAGATCTCCAGCCAGGAGCATCT 130
    |||
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
    |||
QY 131 CCAGCCAGGAGCATCTCCAGCTGGAGACATCTCCAGCCGGGAGATCTCCAGCCAGGAGCATCT 190
    |||
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
    |||
QY 191 CCAGCTGGTACCTCCAGGCGGGGAGCATCTCCAGCCGGGAGATCTCCAGCCAGGAGCATCT 250
    |||
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
    |||
QY 251 CCAGCCGGGAGCATCTCCAGCTGGAGACATCTCCAGCTCCATCCAGGAGGAGTCACTCA 310
    |||
Db 81 ProAlaArgAlaSerProAlaGlnAlaSerLeuSerLeuSerLeuSerLeuSerLeuSer 100
    |||
QY 311 TCCGCGAGGTGAGCTCTGGAGACATCTCCCAACCAAGATGATCTTTAGAGCAACA 370
    |||
Db 101 SerAlaArgSerAlaSerAlaThrThrSerProThrArgValTyrLeuValArgAlaThr 120
    |||
QY 371 CCGTGGGGGCTGATCCAGTCCGATCCATCTCCGCGAGGTCCAGGACGAGCAAGGGGCC 430
    |||
Db 121 ProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrAla 140
    |||
QY 431 ACCAGGAGAGCCAGGAGTACGAGGCTGCCAAGTTCACTGGCGGAGGCGAGCAAGACAG 490
    |||
Db 141 ThrArgGlnSerProGlyThrSerLeuProIlePheThrTrpArgGlyGlnIleGln 160
    |||
QY 491 CTACCGCTCATCGGGTGGTGGTCTCTCTCTCATTTGCTGGTGGTGGTGGTGGTGGTGGT 550
    |||
Db 161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValAlaSerLeuIleIleLeu 180
    |||
QY 551 TTCCAGTTTGGGAGGGGCCACAGAGGATCAGTACAAAGAGAGAGAGAGAGAGTGTGCC 610
    |||
Db 181 PheGlnThrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
    |||
QY 611 AAGCAGCCTGTTCCGCTGAGAGGGGAGTGGAGTCAAGCTGAAGAGTGAAGAGTGGAGC 670
    |||
Db 201 LysHisIleAlaValArgCysAspGlyValValAspCysIleLeuIleSerAlaAspGly 220
    |||
QY 671 TGGGTGAGTTTGAAGTGGAGACAGTCTTGTAAATCTAATCTGGAGTCTCCCATCAG 730
    |||
Db 221 CysValAlaArgPheAspTrpAspIleSerLeuLeuIleTyrSerGlySerSerHisGln 240
    |||
QY 731 TGGGTCCCATCTGTAGAGCAAGTGAAGTCACTCCATAGAGAGAGAGAGAGTGTCCAGAG 790
    |||
Db 241 TrpLeuProIleCysSerSerAspThrAspAspSerTyrSerGlyGlnGlnGlnGlnGln 260
    |||
QY 791 CTGGGTTTCCAGAGTGTCAACCGAGCAACAGAGGTTGGCCACAGAGATTTTCCACAGAC 850
    |||
Db 261 LeuGlyIleGlnSerAlaHisArgThrThrGlyValAlaHisGlyAspPheAlaAsnSer 280
    |||
```

```
QY 851 TTCTCATCTTGATATACAACTCCACCATCCAGAGAAAGCTCCACAGTGTGATGCTT 910
    |||
Db 281 PheSerIleLeuArgTyrIleAsnSerThrIleGlnIleSerLeuHisArgSerGlyCysPro 300
    |||
QY 911 TCCAGCCGATATATCTCCCTCCAGTGTCCCACTGGGAGCTGAGGCCATGACCGGCGG 970
    |||
Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
    |||
QY 971 ATCTGGAGAGGGGCGCTGGCTCCGATAGCAAGTGGCTTTGGCAATGATCTGCACTTC 1030
    |||
Db 321 ILeValGlyGlyAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe 340
    |||
QY 1031 GGCACCAACCAATCTGTGAGGAGCAAGCTATGAGCGCCAGTGGGAGTCACTGCGCGCC 1090
    |||
Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
    |||
QY 1091 CACTGCTTCTTGTGACCCCGGAGAAAGTCTCTGAGGGCTGGAAGGTGAACGGCGGACC 1150
    |||
Db 361 HisCysPhePheValThrArgGlyIleValLeuGlnGlyTyrIleValTyrAlaGlyThr 380
    |||
QY 1151 AGCAACCTGACCAAGTTGCTTGGAGGAGCCTCCATTTGCCAGATCATCAACAGCAAT 1210
    |||
Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
    |||
QY 1211 TACACCGATGAGGAGGACGACTATGACATGCGCTCATGGCGGTGTCAGAGCCCTGACC 1270
    |||
Db 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetCysLeuSerLysProLeuThr 420
    |||
QY 1271 CTGTCCGCTCACAATCCAGCCTGCTGCTCCCAATGATGAGAGACATTTAGCTCAAT 1330
    |||
Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
    |||
QY 1331 GAGACCTGCTGATTCACAGGCTTTGGCAGACCAAGGAGACAGATGACAGACATCCCCC 1390
    |||
Db 441 GluThrCysThrIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
    |||
QY 1391 TTCTCCGGGAGGTGACAGTCAATCATGATGATCAAGAAAGCAATGACTCTTGTC 1450
    |||
Db 461 PheLeuArgGlyValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal 480
    |||
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTGCTGGGAGCTTCGTGGGGGAGAGAC 1510
    |||
Db 481 TyrAspSerTyrIleuThrProArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
    |||
QY 1511 TCTTCCAGGAGACAGCGGGGGGCTCTTGTCTGTGAGCAAGCAACCGCTGTACTCTG 1570
    |||
Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
    |||
QY 1571 GCAGGTGTCAACGCTGGGAGCAGAGCTGTGGCCAGAGAAACAACCTGTGTGTACACC 1630
    |||
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
    |||
QY 1631 AAGTGAAGAGTCTTCCCTGATTTACAGCAAGTGAAGAGAGCGAGTGCATTCAGA 1690
    |||
Db 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetGlnSerGlyValArgPheArg 560
    |||
QY 1691 AAATCC 1696
    |||
Db 561 LysSer 562
    |||
```

RESULT 3
US-10-806-370-12
; Sequence 12, Application US/10806370
; Publication No. US20040209327A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; APPLICANT: Gedrich, Richard
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/10/806,370
; CURRENT FILING DATE: 2004-03-23
; PRIOR APPLICATION NUMBER: US/09/879,792

```

; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-806-370-12

Alignment Scores:
Pred. No.: 1,67e-174 Length: 562
Score: 2999.00 Matches: 562
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-806-370-12 (1-562)
QY 11 ATGGAGAGGAGCAGCAGCGGAGATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
DB 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGAGCATCTCAGCTGGAGCACCTCCAGCGCGGAGCATCTCCAGCCAGGCATCT 130
DB 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGAGCATCTCCAGCTGGAGCACCTCCGCGCGGAGCATCTCCAGCCAGGCATCT 190
DB 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGGTACACCTCCAGCGCGGAGCATCTCCAGCGCGGAGCATCTCCAGCCAGGCATCT 250
DB 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
QY 251 CCAGCCCGGAGCATCTCGGCTCTGGCATCACTTCCAGCTCTGATCCGCGAGGTCATCA 310
DB 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerLeuSerSerSerGlyArgSerSer 100
QY 311 TCCGCCAGGTCAGCTCGGTGACAACTCCCAACCAAGATGTACTTTGTTAGACAA 370
DB 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAGTGGGGGCTGTAACCATCCGATCATCTCCGCCAGGTCAGGACGAGCAACGAGGCC 430
DB 121 ProValGlyAlaValProAlaArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
QY 431 ACCAGGAGAGCCAGGATGAGAGCTGCCCAAGTTCACTGGCGGAGGCGCAGAAACAG 490
DB 141 ThrArgGlnSerProGlyThrThrSerLeuProLysPheThrArgGlnGlyGln 160
QY 491 CTACCGCTCATCGGCGCTGCTCTCTCTCATTCCTGCTGGTGTGCTTCATCATCTC 550
DB 161 LeuProLeuLeuGlyCysValLeuLeuLeuLeuAlaLeuValValSerLeuLeuLeu 180
QY 551 TTCCAGTTCTGGGAGGCGCACACAGAGGATGACAGTACAAAGAGCAGAGGAGAGCTGCC 610
DB 181 PheGlnPheThrArgGlnGlyAlaThrGlyLeuArgTyrLysGlnGlnGlnGlnGlnGln 200
QY 611 AAGCAGCCTGTTCGCTGTGACGAGGATGTGAGTCAAGCTGAAGAGTGAAGAGTGGC 670
DB 201 LysHisLeuAlaValArgCysAspArgValValAlaAspCysLysLeuLysSerArgGlnGly 220
QY 671 TGCCTGAGGTTTGACTGGAGCAAGTCTCTGCTTAAATCTTACTCTGGGCTCTCCATCAG 730

```

```

DB 221 CysValArgPheAspTrpAspLysSerLeuLeuLysIleTyrSerGlySerSerHisGln 240
QY 731 TGGCTTCCCATCTGTACAGCAACTGGATGACTCCACTACAGAGAGACTGGCAGAG 790
DB 241 TrpLeuProLysCysSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 260
QY 791 CTGGGTTTGCAGAGTGTCAACCGGACCAACGAGGTTGCCACAGAGATTTTCCACAGC 850
DB 261 LeuGlyPheGlnSerAlaHisArgThrThrGlnValAlaHisArgAspPheAlaAsnSer 280
QY 851 TTTCATCTTTGAGATACACTCCACCATCCAGAAAGCCTCCACAGTGTGAATGCCCT 910
DB 281 PheSerLeuLeuArgTyrAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCysPro 300
QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCACATGCGGAGCTGAGGAGCATGACCGGCGG 970
DB 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGGGGCGCTGCGCTCGATAGCAGTGGCTTGGCAAGTGAAGTGTGACATTC 1030
DB 321 IleValGlyGlyAlaLeuAlaSerAspSerLysTrpTrpGlnValSerLeuHisPhe 340
QY 1031 GGCACCAACCCACATCTGTGAGGACGCTCATGAGCGCCAGTGGGTGTCACTGCCGCC 1090
DB 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
QY 1091 CACTGCTTCTTGTGACCCGGGAGAGAGTCTGAGAGGCTGGAAGGTTAGCGCGGAGCC 1150
DB 361 HisCysPhePheValThrArgGlnLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
QY 1151 AGCAACCTGACAGATTGCTGAGGACGCTCCATTGCCAGATCATCATCAACAGCAAT 1210
DB 381 SerAsnLeuHisGlnLeuProGlyAlaAlaSerIleAlaGlyIleIleLeuAsnSerAsn 400
QY 1211 TACACGATAGAGAGACGACTATGACATGCGCTCATGCGGCTGTCCAAAGCCCTGACC 1270
DB 401 TyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArgLeuSerLysProLeuThr 420
QY 1271 CTGTCCGCTCACATCCACCTGCTGCTGCCATCGCATGATGACAGACCTTAGCTCAAT 1330
DB 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
QY 1331 GAGACCTGTGATCAACAGCTTTGGCAAGACGAGGAGACAGATGACAAAGATCCGCC 1390
DB 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
QY 1391 TTCTCCGGGAGGTGACAGTCAATTCATGACTTCAAGAAATGCAATGACTTGTGTC 1450
DB 461 PheLeuArgGlnValGlnValAsnLeuIleAspPheLysCysAsnAspTyrLeuVal 480
QY 1451 TATGACAGTAACTTACCCCAAGATGATGTGTGCTGGGAGCCTGCTGGGCGCAGAGAC 1510
DB 481 TyrAspSerTyrLeuThrProAlaGlnMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
QY 1511 TCCTGCCAGGAGACAGCGGGGGGCTTTGTCTGTGAGCAGAAACAACGCTGTGACTGC 1570
DB 501 SerCysGlnGlnLysAspSerGlyLysProLeuValCysGlnGlnAsnAsnArgTrpTyrLeu 520
QY 1571 GCAAGTGTCAACGCTGGGCGACAGGCTGTGGCCAGAGAAACAACTGGTGTGTACACC 1630
DB 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
QY 1631 AAAGTACAGAAATCTTCCCTCGATTTTACAGCAAGATGAGAGCGAGGTGCATTCAGA 1690
DB 541 LysValThrGlnValLeuProTrpIleTyrSerLysMetCysLeuSerGlnValArgPheArg 560
QY 1691 AATCC 1696
DB 561 LysSer 562

```

RESULT 4
US-10-156-214A-24
Sequence 24, Application US/10156214A

Publication No. US20040001801A1
GENERAL INFORMATION:
APPLICANT: Edwin L. Madison
APPLICANT: Joseph Edward Sempke
APPLICANT: George P. Vlasuk
APPLICANT: Scott Jeffrey Kemp
APPLICANT: Mallareddy Komandla
APPLICANT: Daniel Vanna Siev
TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
FILE OF INVENTION: Thereof
FILE REFERENCE: 24745-1611
CURRENT APPLICATION NUMBER: US/10/156,214A
CURRENT FILING DATE: 2002-05-23
NUMBER OF SEQ. ID NOS: 611
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 24
LENGTH: 562
TYPE: PRT
ORGANISM: homo sapien
FEATURE:
NAME/KEY: protease domain of endothelialase 2
LOCATION: (321)..(562)
US-10-156-214A-24

Alignment Scores:

Pred. No.:	5,14e-174	Length:	562
Score:	2991.00	Matches:	561
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	1
Query Match:	89.7%	Indels:	0
DB:	4	Gaps:	0

US-10-806-370-11 (1-1748) x US-10-156-214A-24 (1-562)

11 ATGAGAGGAGCAAGCAGGAGATGATCTTCAGACAAGAACCTTCAGCTGAGCATCT 70
1 MeGluNrgApsrHrHsglyAsnAlaSerProAlaHrghrProSerAlaGlyAlaSer 20
71 CAGGCCAGGATCTCAGCTGGAGACCTTCAGGCCGGGAGCATCTCAGGCCAGGATCT 130
21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
131 CAGGCCAGGATCTCAGCTGGAGACCTTCAGGCCGGGAGCATCTCAGGCCAGGATCT 190
41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
191 CAGCTGATACACTCCAGGCCGGGAGCATCTTCAGGCCGGGAGCATCTCAGGCCAGGATCT 250
61 ProAlaGlyHrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
251 CAGGCCGGGAGCATCTCCGGCTCTGGCATCTTCAGGTCCTTCATCCGGAGGATCTCA 310
81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
311 TCCGCCAGGTCAGCTCGGTGGAGAACCTCCCAACAGAGTGCCTTTAGAGCAACA 370
101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrlleuValAlaGlaThr 120
371 CAGTGGGGCTGTACCATCTCCGATCTTCCTCCAGCTCAGGACAGCAACAGAGGCC 430
121 ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
431 ACCAGGAGAGCCAGGATGAGAGCTTCGCCAAGTTCACTGGCGGAGAGGCCAGAGCAG 490
141 ThrArgGlnSerProGlyThrSerLeuProLeuSerThrTrpArgGlnGlyGlnLeu 160
491 CTACCGCTCACTCGGGTGGTGGCTCTCTCCTCATTCCTCCGAGGTTTGCTCATCATCTC 550
161 LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerLeuIleIleu 180
551 TTCGAGTTTGGAGGGCCACACAGGAGATCAGTACAGGACAGAGAGAGAGCTGTCCC 610
181 PheGlnHrghrPrpGlnGlyHsrThrGlyIleArgTyrlleGlnGlnArgGlnSerCysPro 200

611 AAGCAAGCTGTGGCTGTGACGGGGTGGTGGACTGCAAGCTGAAGAGTGAAGAGCTGGGC 670
201 LyhIleAlaValAlaArgCysApsrGlyValAlaAspCysLeuSerLeuSerApsrGlyLeuGly 220
671 TGGGTGAGTTTGAAGTGAAGCAAGTCTCTGCTTTAAATCTACTCTGGTCTCCATCAG 730
221 CysValaArgPheApsrTrpApsrlySerLeuLeuValIleTySerGlySerSerHsGln 240
731 TGGCTTCCATCTGTAGACCAACTGGAATGACTCTTACTACAGAAAGACTGACACAG 790
241 TrpLeuProIleCysSerSerApsrTrpApsrSerTySerGlyuSerTySerGlnGln 260
791 CTGGCTTTCAGAGTGTGTCACCGGACACCGAGTGGCCCAAGGATTTTGGCAAGC 850
261 LeuGlyPheGlnSerAlaHsrApsrThrGlyValAlaHsrApsrPheAlaApsr 280
851 TTCATCTTGAATCAACTCCACCATCCAGAAAGCTTCACAGAGTGTGAATGSCCT 910
281 PheSerIleLeuArgTyrlaAspSerThrIleGlnGlnSerLeuHsrArgSerGlyCysPro 300
911 TCCAGCGGTATATCTCCCTCCAGTGTCCCACTGCGGACTGAGGSCATGACCGGCGG 970
301 SerGlnArgTyrlleSerLeuGlnCysSerHsCysGlyLeuArgAlaMetThrGlyArg 320
971 ATGCTGGAGAGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAAGTGTGCACTTC 1030
321 IleValGlyIleAlaLeuAlaSerApsrSerTyTrpProTrpGlnValSerLeuHsrPhe 340
1031 GGCACCAACCAATCTGTGAGGACCGCTCATTTGACGCCCGAGTGGGTCTCACTGCGGCC 1090
341 GlyThrThrHsrIleCysGlyGlyThrLeuIleApsrAlaGlnTrpValLeuThrAlaAla 360
1091 CACTGCTTCTTGTGACCCGGGAGAAAGTCTGAGGCTGGAAGTGAACGGCGGACCC 1150
361 HsCysPhePheValThrArgGlyuValLeuGlyGlyTrpTyValTyrlaGlyThr 380
1151 AGCAACTGACCAAGTTGCTGAGGACGCTCCATTCGCGAGATCATCATCAACAGCAAT 1210
381 SerAsnLeuHsrGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleApsrApsr 400
1211 TACACGATGAGGAGGACGACTATGACATGCGCTCATGCGGCTGTCCAGGCCCTGACC 1270
401 TyrlThrApsrIleGlnApsrTyrlApsrIleAlaLeuMetArgLeuSerTyrlProLeuThr 420
1271 CTGTCCGCTCAATCCACCTGCTGCTCCCATGATGAGACAGACTTTAGCTCAAT 1330
421 LeuSerAlaHsrIleHsrProAlaCysLeuProMetHsGlyGlnThrPheSerLeuApsr 440
1331 GAGACCTGCTGATCAACAGCTTTGGCAAGACCAAGGAGACAGATGACAGACATCCCCC 1390
441 GlnThrCysApsrIleThrGlyPheGlyTyThrArgGlnThrApsrApsrTyrlThrSerPro 460
1391 TTCCTCCGGAGGTCAGGTCATCATGATGACTTCAAGAAATGCAATGACTACTTGCTG 1450
461 PheLeuArgGlnValGlnValApsrLeuIleApsrPheLeuTyrlCysApsrApsrTyrlleuVal 480
1451 TATGACATTAACCTTACCCCAAGATGATGTGTCTGGGAGCTTCGTGGGGCCAGAGAC 1510
481 TyrlApsrSerTyrlleuThrProArgMetCysAlaGlyApsrLeuArgGlyGlyArgApsr 500
1511 TCTGTCAGGAGGACACAGCGGGGGCTCTGTGTGTGACGAGCAACCAACCTGTGACTCTG 1570
501 SerCysGlnGlyApsrSerGlyGlyProLeuValCysGlnGlnApsrApsrApsrTyrlleu 520
1571 GCAGTGTCAACAGCTGGGCAACAGCTGTGAGCCAGAGAAACCAACCTGTGTGTACACC 1630
521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgApsrApsrProGlyValTyrlThr 540
1631 AAAGTGAACAGAGTTCTTCCCTGATTTTACAGCAAGATGAGAGCGAGTGCATTCAGA 1690
541 TyrlValThrIleValLeuProTrpIleTyrlSerTyrlMetGlnSerGlnValArgPheIle 560

QY 1691 AATCC 1696
 Db 561 LysSer 562

RESULT 5

US-10-916-758-94
 ; Sequence 94, Application US/10916758
 ; Publication No. US20050180977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nixon, Andrew
 ; APPLICANT: Madison, Edwin L.
 ; TITLE OF INVENTION: ENDOTHELINASE-2 LIGANDS
 ; FILE REFERENCE: 10280-065001
 ; CURRENT APPLICATION NUMBER: US/10/916,758
 ; PRIOR FILING DATE: 2004-08-12
 ; PRIOR APPLICATION NUMBER: US 60/520,164
 ; PRIOR FILING DATE: 2003-11-14
 ; PRIOR APPLICATION NUMBER: US 60/495,005
 ; NUMBER OF SEQ ID NOS: 113
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 94
 ; LENGTH: 562
 ; TYPE: PR
 ; ORGANISM: Homo sapiens
 US-10-916-758-94

Alignment Scores:

Pred. No.: 5,14e-174 Length: 562
 Score: 2991.00 Matches: 561
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 89.7% Indels: 0
 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-916-758-94 (1-562)

QY 11 ATGAGAGAGGACAGCCAGGGAATGCATCTTCAGCAAGAACCTTCAGCTGAGCATCT 70
 Db 1 MetGlnArgSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
 QY 71 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGCCGGGATCTCCAGCCAGGATCT 130
 Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
 QY 131 CCAGCCAGGATCTCCAGCTGGAGACACTCCAGCCGGGATCTCCAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
 QY 191 CCAGCTGATACACTCCAGCCGGGATCTTCAGCCGGGATCTTCAGCCAGGATCT 250
 Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
 QY 251 CCAGCCGGGATCTCCAGCTGGAGACACTTCAGCTGATCTTCAGCTGATCTTCAGCTGATCT 310
 Db 81 ProAlaArgAlaSerProAlaGlnAlaSerLeuSerLeuSerSerSerGlyArgSerSer 100
 QY 311 TCCGCCAGGATCGAGCTCGGTGACAACTCCCAAGTTCACTTCGCGGAGGAGGAGCAAGAG 370
 Db 101 SerLeuArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
 QY 371 CCAGTGGGGGTGTACCATTCGATTCATCTCTGCCAGTGCAGACCAAGCAAGGAGCC 430
 Db 121 ProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
 QY 431 ACCAGGAGAGACCCAGGATACAGAGCTGCCCAAGTTCACTTCGCGGAGGAGGAGCAAGAG 490
 Db 141 ThrArgLeuSerProGlyThrSerLeuProLysPheThrTrpArgGlnGlnAlaGln 160
 QY 491 CTACCGCATCGAGGTGCTCTCTCATTCGCGCTGAGTGTTCGCTCATCATCTCTC 550
 Db 161 LeuProLeuLeuGlyCysValLeuLeuLeuLeuLeuLeuValSerLeuLeuLeuLeu 180

QY 551 TTCAGATTCTGGACGGGCCACACAGGATCAGGTACAGAGACAGAGGAGGAGCTGTCC 610
 Db 181 PheGlnPheTrpGlnGlyHisThrGlyIleArgTyrLeuGlnArgGlnSerCysPro 200
 QY 611 AAGCAAGCTGTTCGCTGTGACCGGGGTGTGAGCTGCAGACTGAAGACTGACAGCTGGC 670
 Db 201 LysHisAlaValArgCysAspGlyValValAspCysLeuValSerAspGlnLeuGly 220
 QY 671 TGCCTGAGTTTGAAGTGGACAAAGTCTCTGCTTAAATCTACTCTGGGCTCTCCATCAG 730
 Db 221 CysValArgPheAspTrpAspLysSerLeuLeuValLeuTyrSerGlySerSerHisGln 240
 QY 731 TGCTTCCATCTGTACAGCACTGGAATGACTCTTACTCAGAGAACTGTCACAG 790
 Db 241 TrpLeuProIleCysSerSerAspTrpAsnAspPheTyrSerGlyThrCysGlnGln 260
 QY 791 CTGGGTTTCAGAGCTGTCACCCGACCAAGAGGTTGCCACAGGATTTTGGCAACAG 850
 Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGluValAlaHisArgAspPheAlaSerSer 280
 QY 851 TTCTCAATCTTGAGATACAACTCCACCATCCAGAGAAAGCTCCACAGATCGAATGCC 910
 Db 281 PheSerIleLeuArgTyrAspSerThrIleGlnGlnSerLeuHisArgSerGluCysPro 300
 QY 911 TCCAGCGGTATATCTCCCTCCAGTGTTCACCTGCGCATGAGGAGCCATGACCGGCG 970
 Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
 QY 971 ATCGTGGAGGGGGGCTGGGCTGGATAGCAAGTGGCTTGGCAAGTGAAGCTTCACATTC 1030
 Db 321 IleValIleGlyAlaLeuAlaSerAspSerTyrTrpTrpGlnValSerLeuHisPhe 340
 QY 1031 GGCACCAACCAATCTGTGAGAGGACCGCTCATGGACCCAGTGGTGTCTACTCGCC 1090
 Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
 QY 1091 CACTGCTTCTTCGTGACCCCGAGAGAGCTCTGAGAGGCTGAAAGGTGTACCGGGCAC 1150
 Db 361 HisCysPhePheValThrArgGlyValLeuGlnIleTyrIleValIleGlyThr 380
 QY 1151 AGCAACCTGACCAAGTTCGCTGGAGGAGCTCCATTCGCCAGATCATCAACAGCAAT 1210
 Db 381 SerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGlnIleIleIleAsnSerAsn 400
 QY 1211 TACACGATGAGAGAGAGCACTATGATGATGACGCTCCATGCGGCTGTCCAAAGCCCTGAC 1270
 Db 401 TyrThrAspGlnGluAspAspTyrAspIleAlaLeuValAspArgLeuSerLysProLeuThr 420
 QY 1271 CTGTCCGCTCATCATCCACCTGCTGCTGCCATGCAATGACAGACCTTTAGCTCAAT 1330
 Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
 QY 1331 GAGACCTGCTGATCAACAGCTTTGGACAGACGAGGAGACAGATGACAAGATCCCCC 1390
 Db 441 GlnThrCysTrpIleThrGlyPheGlyLysThrArgGlnThrAspAspLysThrSerPro 460
 QY 1391 TTCCTCCGAGAGGTGACGATCAATCTCATGACTTCAAGAAATGCAATGACTACTGTGC 1450
 Db 461 PheLeuArgIleValGlnValAlaLeuLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
 QY 1451 TATGACAGTTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCGTGGGGGAGAGAG 1510
 Db 481 TyrAspSerTyrLeuThrTrpArgMetMetCysAlaGlyAspLeuArgGlyGlyArgAsp 500
 QY 1511 TCCGTGCAAGGAGACAGCGGGGGGCTCTGTCTGTGAGCAAGAACCGGTGTACTGCTG 1570
 Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnAsnAspTrpTyrLeu 520
 QY 1571 GCAAGTGTCAACAGCTGGGACACAGGCTGTGAGCCAGAGAAACAACTGTGTGTACACC 1630
 Db 521 AlaGlyValHisSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
 QY 1631 AAAGTGAACAGAGTTCTTCCTCGATTTACAGCAAGATGAGAGCGAGGTGCGATTGAGA 1690

Db 541 LysValThrGluValLeuProTPrIleTyrSerIleMetGluSerGluValArgPheIle 560
QY 1691 AATCC 1696
Db 561 LysSer 562

RESULT 6
US-10-156-214A-26
; Sequence 26, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Semple
; APPLICANT: George P. Vlasak
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156, 214A
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: protease domain
; LOCATION: (321)..(688)
US-10-156-214A-26

Alignment Scores:
Pred. No.: 2,69e-172 Length: 688
Score: 2963.00 Matches: 554
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 88.9% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-156-214A-26 (1-688)

QY 11 ATGAGAGGAGACAGCCAGGAGATGATCTTCACCAAGAACACTTCACTGAGCATCT 70
Db 1 MetGluArgSerSerIleGlyAsnAlaSerProIleArgThrProSerAlaGlyAlaSer 20
QY 71 CCAAGCCAGGATTCAGCTGGAGGACCTCCAGCCGGGATCTCCAGCCAGGATCT 130
Db 21 ProIleGlnAlaSerProIleGlyThrProProGlyArgAlaSerProIleGlnAlaSer 40
QY 131 CCAAGCCAGGATCTCCAGCTGGAGGACCTCCAGCCGGGATCTCCAGCCAGGATCT 190
Db 41 ProIleGlnAlaSerProIleGlyThrProProGlyArgAlaSerProIleGlnAlaSer 60
QY 191 CCAAGCTGATACCTCCAGGCGGAGCATCTCCAGGCGGAGCATCTCCAGCCAGGATCT 250
Db 61 ProIleGlyThrProProGlyArgAlaSerProIleGlyArgAlaSerProIleGlnAlaSer 80
QY 251 CCAAGCCGGGATCTCCGGGCTGGGATCTTCACGATCTCTGATCCGGGAGGATCA 310
Db 81 ProIleArgAlaSerProIleAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer 100
QY 311 TCCGCGAGGTCAGCTCGGTGAGCAACCTCCCAACCAAGATGTAACCTTTAGAGCAACA 370
Db 101 SerIleArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
QY 371 CCAAGTGGGGCTGTACCATCTCCATCTCTCTGCGAGGTAGCAACCAAGCAAGGAGCC 430
Db 121 ProValGlyAlaValProIleArgSerSerProIleArgSerAlaProIleAlaThrArgAla 140
QY 431 ACCAGGAGAGACCCAGGTAGAGAGCTGCGCAAGTTACCTGCGGAGGAGGCGAGAGAG 490

Db 141 ThrArgIleSerProGlyThrSerIleuProIlePheThrIleArgGluGlnIleuGln 160
QY 491 CTACCGCTCATCGGAGCGGTGCTCTCTCATTTGCCCTGGTGGTTGCTCATATCTCTC 550
Db 161 LeuProIleuIleGlyCysValLeuIleuIleAlaLeuValIleSerIleIleIleu 180
QY 551 TTCAGTTCTGGAGGCGCACAGAGGATGAGTACAGAGGAGAGAGAGAGAGAGAGAGAG 610
Db 181 PheGlnPheThrGlnGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 200
QY 611 AACGACGCTTTGCTGTGACAGGAGGAGGAGTGTGAGTGTGACAGTGTGAGTGTGAGTGTG 670
Db 201 LysIleAlaValAlaArgCysArgGlyValAlaAspCysIleuIleuIleuIleuIleuGly 220
QY 671 TGGGTAGGTTTGAAGTGTGAGCAAGTCTGTGTTAAATTAATCTGTGGTCTCCATCAG 730
Db 221 CysValArgPheArgPheArgPheArgPheArgPheArgPheArgPheArgPheArgPhe 240
QY 731 TGGCTTCCCATCTGTGACAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 790
Db 241 TyrLeuProIleCysSerSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 260
QY 791 CTGGGTTTCCAGAGTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 850
Db 261 LeuGlyPheGluSerAlaIleIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 280
QY 851 TTCTCAATCTTGAGATCAATCCACCATCCAGAAAGCTCCACAGTGTGAATGCTCT 910
Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnIleuSerIleuIleuIleuIleuIleuIleu 300
QY 911 TCCGAGGATATATCTCCCTCCAGTCTTCCACATGCGGAGTGTGAGGAGGAGGAGGAGG 970
Db 301 SerGlnArgTyrIleSerIleuGlnCysSerIleuGlyIleuArgAlaMetThrGlyArg 320
QY 971 ATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1030
Db 321 IleValGlyIleAlaLeuAlaSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 340
QY 1031 GGCACACCCACATCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090
Db 341 GlyThrThrIleIleCysGlyGlyThrLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 360
QY 1091 CACTGCTTTCTTGTGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1150
Db 361 HisCysPhePheValThrArgGluIleuValLeuGlnIleuIleuIleuIleuIleuIleuIleu 380
QY 1151 AGCAAGCTGACAGGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1210
Db 381 SerAsnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 400
QY 1211 TACACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1270
Db 401 TyrThrArgGluGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
QY 1271 CTGTCCGCTCATCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
Db 421 LeuSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 440
QY 1331 GAGACGCTGATACAGAGCTTTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1390
Db 441 GluThrCysTrpIleThrGlyPheGlyIleuIleuIleuIleuIleuIleuIleuIleuIleu 460
QY 1391 TTCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1450
Db 461 PheLeuArgGluValGlnValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 480
QY 1451 TATGACGATTAACCTTACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1510
Db 481 TyrAspSerTyrLeuThrProIleArgMetCysAlaIleuIleuIleuIleuIleuIleuIleu 500
QY 1511 TCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1570

Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnArgTrpTyrLeu 520
Qy 1571 GCAGGTGTACACGAGCTGGGGGACAGGCTGTGGCCAGAGAAACAACCTGGTGTACAC 1630
Db 521 AlaGlyValThrSerTrpGlyThrGlyCysGlyGlnArgAsnLysProGlyValTyrThr 540
Qy 1631 AAAGTGAAGAGAGTTCTTCCCTGGATTTTACGCAAGATGAGAGCCGAGTCCGA 1684
Db 541 LysValThrGluValLeuProTrpTyrLeuTyrSerLysMetGluAsnArgAlaGln 558
RESULT 7
US-10-916-758-2
; Sequence 2, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELINASE-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-2
Alignment Scores:
Pred. No.: 2,696-172 Length: 688
Score: 2963.00 Matches: 554
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 88.9% Indels: 0
DB: 5 Gaps: 0
US-10-806-370-11 (1-1748) x US-10-916-758-2 (1-688)
Qy 11 ATGAGAGAGGACACGACCGGGAATGATCTTCAGCAAGAACACCTTCAGTGAACATCT 70
Db 1 MetGluArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
Qy 71 CCAGCCAGGATCTTCAGCTGGGACACCTTCAGGCGGGGATCTTCAGCCAGGACATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 40
Qy 131 CCAGCCAGGATCTTCAGCTGGGACACCTTCGCGGGCGGGATCTTCAGCCAGGACATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer 60
Qy 191 CCAAGTGTATACCTTCAGGCGGGGATCTTCAGGCGGGGATCTTCAGCCAGGACATCT 250
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer 80
Qy 251 CCAGCCCGGGGATCTTCGCGCTTCGACATCACTTTTCAGGTCTTCATCCGCGAGGTCACTCA 310
Db 81 ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerSerGlyArgSerSer 100
Qy 311 TCCGCGAGGTGAGCTCGGTGACAACTCCCAACAGAGTGAAGTTCCTGTAGAGACA 370
Db 101 SerAlaArgSerAlaSerValThrThrSerProThrArgValTyrLeuValArgAlaThr 120
Qy 371 CCAAGTGGGGGTGTATACCATCGATCATCTCTCGCAGGTGAGACACAGAACAGGGGCC 430
Db 121 ProValGlyAlaValProLeuArgSerSerProAlaArgSerAlaProAlaThrArgAla 140
Qy 431 ACCGAGGAGAGCCACAGTACAGAGCTGCCCAAGTTCACTTGGCGGAGGGCCAGAGACG 490

Db 141 ThrArgLysSerProGlyThrSerLeuProLysBheThrTrpArgGlnGlyGlnLysGln 160
Qy 491 CTACCGGTGATCGGGGTGCGTGTCTCTCTCATTCGCTCGGTGGTGTTCATCATCTCTC 550
Db 161 LeuProLeuLeuGlyCysValLeuLeuLeuLeuAlaLeuValAlaSerLeuLeuLeu 180
Qy 551 TTCAGTTCTGGACAGGGCCACACAGGATGACGTACAGAGACAGAGAGAGCTGTCC 610
Db 181 PheGlnThrProGlnGlyHisThrGlyLeuArgTyrLysGlnGlnArgLysSerCysPro 200
Qy 611 AAGACGCTGTTCGCTGTGACCGGGGTGTCGTCGACGTCGAAAGTTCAGACTGGCC 670
Db 201 LysHisAlaValArgCysAspGlyValValAspCysLysLeuLysSerAspGlyLeuGly 220
Qy 671 TGCGTGAGGTTTGACTGGGACAAAGTCTCTGCTTAAATCTACTCTGGGCTCTCCATCAG 730
Db 221 CysValArgPheAspTrpAspLysSerLeuLeuLysLeuTyrSerGlySerSerHisGln 240
Qy 731 TGGCTTCCATCTGTACACCACTGGAATGACTCTTACTCAGAGAACCTGCCACAG 790
Db 241 TrpLeuProIleCysSerSerAsnTrpAsnAspSerTyrSerGlyLysThrCysGlnGln 260
Qy 791 CTGGGTTTCGAGAGTGTCAACCGGACAAACGAGGTTGCCCCAGAGGATTTTCCCAAGC 850
Db 261 LeuGlyPheGlnSerAlaHisArgThrThrGluValAlaHisArgAspPheAlaAsnSer 280
Qy 851 TTCTCATCTTGAGATACAACTCCACCATCCAGAGAACCTCCACAGGTCTGAATGCCCT 910
Db 281 PheSerIleLeuArgTyrAsnSerThrIleGlnLysSerLeuHisArgSerGlyCysPro 300
Qy 911 TCCAGCGGTATATCTCCCTCCAGTGTCCCACTGCGGACTGAGGGCATGACCGGGCGG 970
Db 301 SerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArgAlaMetThrGlyArg 320
Qy 971 ATCTGGGAGGGGCGCTCGGCTCGGATAGCAAGTGGCTTGGCAAGGAGTTCGACATTC 1030
Db 321 IleValGlyValAlaLeuAlaSerAspSerLysTrpProTrpGlnValSerLeuHisPhe 340
Qy 1031 GGCACACCAACATCTGTGAGAGGACCGCTCATTTGACCCGAGTGGGTCTCACTCGCCGC 1090
Db 341 GlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrpValLeuThrAlaAla 360
Qy 1091 CACTGCTTCTTCTGTGACCCGGGAGAGAGTCTTGAGGGCTGGAAGGTGTACCGGGCAC 1150
Db 361 HisCysPhePheValThrArgGlyLysValLeuGlnGlyTrpLysValTyrAlaGlyThr 380
Qy 1151 AGCAACCTGACACCGTTCGCTGAGGACGCTCCATTCGAGATCATCATCAACAGCAAT 1210
Db 381 SerAsnLeuHisGlnLeuProGluAlaAlaSerIleAlaGluIleIleIleAsnSerAsn 400
Qy 1211 TACACGATGAGAGAGACGACTATGACATGCGCTTCATGCGGCTGTCCAAAGCCCTGACC 1270
Db 401 TyrThrAspGluGluAspAspTyrAspIleAlaLeuMetArgLysSerLysProLeuThr 420
Qy 1271 CTGTCCCTCAATCCACCTCTTGCCTCCCATGATGACAGACCTTTAGCTCAAT 1330
Db 421 LeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGlnThrPheSerLeuAsn 440
Qy 1331 GAGACCTGCGAGTACAGAGCTTTGGAAGACACGAGGAGACAGATGACAGATCCCCC 1390
Db 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAspAspLysTrpSerPro 460
Qy 1391 TTCCTCCGAGAGGTGACGATCAATCTCATGACTTCAAGAAATGCAATGACTTGTGTC 1450
Db 461 PheLeuArgGluValGlnValAlaLeuLeuIleAspPheLysLysCysAsnAspTyrLeuVal 480
Qy 1451 TATGACATTAACCTTACCCCAAGATGATGTGTGCTGGGACCTTGGTGGGGCAGAGAC 1510
Db 481 TyrAspSerTyrLeuThrProArgMetMetCysAlaGlyLysAspLeuArgGlyGlyArgAsp 500
Qy 1511 TCCGCGCAGGAGACAGCGGGGGGCTCTGTCTGTGAGAGAAACAACCGCTGTACTCTG 1570
Db 501 SerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsnArgTrpTyrLeu 520

```

QY      1571   GCAGGTGCACCACTGGGCGCAGCGGTCGCCAGGAAGAAACAACCTGGTGTCCACC    1630
Db      521    AAcgtgaaatmrsetrrtgelYthGelCyseLylmrgshnlyspGoglyValTYrThn    540

QY      1631   AAAGTGACAGAAGTCTTCCTCGATTATTAACGACAAGATGAGAGCGCATGCCA    1684
Db      541    LysValThrGUvalLeuPTrOTPrILeTySerLysetGluMetGluAsnAgaLAaGln    558

RESULT 8
US-10-353-690-100
; Sequence 100, Application US/10353690
; Publication No. US20030215840A1
GENERAL INFORMATION:
APPLICANT: Logan, Thomas Joseph
APPLICANT: Chun, Miyoung
APPLICANT: Galvin, Katherine M.
APPLICANT: Healy, Aileen
APPLICANT: Acton, Susan L.
APPLICANT: Donoghue, Mary
APPLICANT: Stagliano, Nancy
APPLICANT: Perodin, Jacqueline
APPLICANT: Rodrique-Way, Amelie
TITLE OF INVENTION: Methods and compositions for treating
TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395
TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720
TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3464, 345, 9252, 9135,
TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
TITLE OF INVENTION: 51130, 19469, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
FILE REFERENCE: MP102-018PIRMONNIM
CURRENT APPLICATION NUMBER: US/10/353,690
CURRENT FILING DATE: 2003-01-29
PRIOR APPLICATION NUMBER: 60/353,224
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 60/364,529
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: 60/373,861
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/376,287
PRIOR FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 60/388,080
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: 60/390,971
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/394,130
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: 60/394,797
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/404,904
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/405,450
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 100
LENGTH: 581
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-353-690-100

Alignment Scores:
Pred. No.: 7.03e-172 Length: 581
Score: 2956.00 Matches: 553
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 88.7% Indels: 0
DB: 4 Gaps: 0

US-10-806-370-11 (1-1748) x US-10-353-690-100 (1-581)

```

QY	1	ATGAGAGGGGACACCCAGGGAAAGCATCTCCAGCAAGAAACACCTTCAGCTGGAGCATCT	70
Db	1	MetGluArgSerHisTSGLYAsnAlaSerProAlaGlnProSerAlaGlyAlaSer	20
QY	71	CCAGCCGAGGATCTCCAGCTGGGACACCTCCAGCCGGGCAATCTCCAGCCAGGCATCT	130
Db	21	ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer	40
QY	131	CCAGCCGAGGATCTCCAGCTGGGACACCTCCGGGCGGGGCAATCTCCAGCCAGGCATCT	190
Db	41	ProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAlaSer	60
QY	191	CCAGCTGGTACACTCCAGGCGGGGCAATCTCCAGCCGGGCAATCTCCAGCCAGGCATCT	250
Db	61	ProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAlaSer	80
QY	251	CCAGCCGGGCAATCTCCGGGCTTGGCATCACTTCCAGGCTCTCATCCGGGAGTCATCA	310
Db	81	ProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSerSerGlyArgSerSer	100
QY	311	TCCGCGAGGATGAGCTCGGTGACAACTCCGCCAACAAGAGTGCCTTGTTAAGCAACA	370
Db	101	SerAlaGSerAlaSerValThrThrSerProThrArgAlaTyrLeuValArgAlaThr	120
QY	371	CCAGTGGGGGTGTATCCATCCGATCATCTCTGCGAGGTACGACCAAGCAACCGAGGCC	430
Db	121	ProValGlyAlaValProIleArgSerSerProAlaArgSerAlaProAlaThrArgAla	140
QY	431	ACCAAGGAGAGCCCAAGGTACGAGCTCCGCCAAGTTCACCTGGCGGAGGGCCAGAACAG	490
Db	141	ThrArgGlnSerProGlyThrSerLeuProLysPheThrTrrArgGlnGlyGlnLysGln	160
QY	491	CTACCGGTCATCGGGGTGCGTGCCTCTCTCATTCGCCCTGGTGGTGTGGTCATCATCTCTC	550
Db	161	LeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValValSerIleIleIleLeu	180
QY	551	TTCCAGTTCTTGGCAGGGGCCACACAGGATCAGGTACAGAGACAGAGAGAGCTGTCCC	610
Db	181	PheGlnPheThrGlnGlnIleHisThrGlyIleArgHisIleLysGlnGlnArgGlnLysCysPro	200
QY	611	AAGGACGCTGTGGCTGTGACCGGGGTGGTGGACGTGCAAGGTCGAAGACTGACAGCTGGGC	670
Db	201	LysHisAlaValAlaGlyCysArgGlyValValLysProLysLysLeuLysSerLysArgGlnLysGly	220
QY	671	TGCGTGAAGTTGACTGGGACAACTCTCTGTTAAATTACTCTGGGCTCTCCCATCAG	730
Db	221	CysValAlaArgPheAspTrrPheLysSerLeuLeuLysIleTyrSerGlySerSerHisGln	240
QY	731	TGGCTTCCCATCTGTAGCAGAACACTGGAAAGACTCTCTACAGAGAAAGACTGCGACAG	790
Db	241	TrrLeuProIleCysSerSerLysAsnTrrLysAsnArgSerTyrSerGlyLysThrCysGlnGln	260
QY	791	CTGGGTTTCCAGATGTCTCACCGGACAAACGAGGTTTCCACAGGATTTTGGCAACAGC	850
Db	261	LeuGlyPheGlnSerAlaHisArgGlyThrThrGlyValAlaHisArgAsnArgPheAlaHisSer	280
QY	851	TTTCTCAATCTTGAGATCAAACTCCACCATCCAGGAAGCTTCCACAGCTGTAATGCTCCT	910
Db	281	PheSerIleLeuAlaGlyLysAsnSerThrIleGlnGlnSerLeuHisArgSerGlnCysArgPro	300
QY	911	TCCAGAGGGATATCTCCCTCCCAAGTGTCCACGTGGGAGCTGAAGGGCAGACCGGGCGG	970
Db	301	SerGlnArgTyrIleSerLeuGlnCysSerHisCysAlaLeuArgAlaMetThrGlyArg	320
QY	971	ATCTGGGAGGGGCGCTGGCCTCGGATAGCAAGTGCCTTGGCAAGTGAAGTTCGACTTC	1030
Db	321	IleValGlyGlyAlaLeuAlaSerAspSerLysTrrProTrrGlnValSerLeuHisArgPhe	340
QY	1031	GGGACCAACCAATCTGTGAGGAGCAACCTATGACGCCAAGTGGGCTCAGTCCGCGCC	1090
Db	341	GlyThrThrHisIleIleCysGlyGlyThrLeuIleAsnArgAlaGlnTrrValLeuThrAlaAla	360

QY 1091 CACTGCTTCTTCTGAGCCCGGAGAGAGTCTCTGAGGCGCTGAGAAAGTGTACGCGGCGAC 1150
DB 361 HieCysephneValThrArgGluValLeuGluGlyTyrPlyValTylAlaGlyThr 380
QY 1151 AGCAACTGACACCGATTGCTGAGGAGCGCTCCATTCGCCGATCATCATCAACAGCAAT 1210
DB 381 SerAenLeuHlaGlnLeuProGluAlaSerIleAlaGluIleIleIleAenSerAen 400
QY 1211 TACACCGATGAGGAGACATGATGACATGCGCCCTCAAGCGGCTGTCCAAACCCCTGAC 1270
DB 401 TyrThrAepGluGluAepApyThrAepIleAlaLeuMetAglLeuSerLysProLeuThr 420
QY 1271 CTGTCGCTCATCATCCCTCTGCTGCTCCCATGATGACAGACCTTATGCTCAAT 1330
DB 421 LeuSerAlaHlaSileHlaProAlaCylLeuProMetHlaGlyGlnThrPheSerLeuAen 440
QY 1331 GAGACCTGCTGATATCAACGCTTTGGCAAGACAGGAGACAGATGACAAACATCCGCC 1390
DB 441 GluThrCysTrpIleThrGlyPheGlyLysThrArgGluThrAepApyLysThrSerPro 460
QY 1391 TTCTCCGAGAGTGCAGATCATCTCATGACTTCAAGAAATGCAATGACTTACTGCTC 1450
DB 461 PheLeuAglValAlaGlnValAenLeuIleAepPheLysCysAenAepTyrLeuVal 480
QY 1451 TATGACAGTTACCTTACCCCAAGATGATGTGTCTGGGACCTTCTGCGGCGAGAGAC 1510
DB 481 TyrAepSerTyrLeuThrProArgMetMetCysAlaGlyAepLeuArgGlyGlyArgAep 500
QY 1511 TCTTCGCGGAGAGACAGCGGGGGCTTGTGTCTGTGAGCAACACCGCTGTACTGT 1570
DB 501 SerCysGlnGlyAepSerGlyGlyProLeuValCysGlnGlnAenAenAenAenAen 520
QY 1571 GCAGGTGTCACACGCTGGGAGACAGGCTGTGGCCAGAGAAACAACTGGTGTACACC 1630
DB 521 AlaGlyValThrSerTyrGlyThrGlyCysGlyGlnAenAenAenAenAenAenAen 540
QY 1631 AAAGTGACAGAGTCTTCTCTGATTTTACAGCAAGATGAG 1672
DB 541 LysValThrGluValLeuProTyrIleTyrSerLysMetGlu 554

RESULT 9
US-10-428-275-120
; Sequence 120, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OR INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428, 275
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 120
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-428-275-120

Alignment Scores:
Pred. No.: 1e-171 Length: 593
Score: 2953.50 Matches: 555
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 0
Query Match: 88.6% Indels: 5
DB: 4 Gaps: 1
US-10-806-370-11 (1-1748) x US-10-428-275-120 (1-593)
QY 8 ACATGAGAGAGGACACCCAGGAAATGATCTTCACCAAGAACACTTCACTGAGCA 67
DB 4 ThrMetGluArgAepSerHlaGlyAenAlaSerProAlaArgThrProSerAlaGlyAla 23
QY 68 TCTCCAGCCAGGATCTCCAGCTGGGACCTCCAGCGGGGATCTCCAGCCAGGCA 127
DB 24 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 43
QY 128 TCTCCAGCCAGGATCTCCAGCTGGGACCTCCAGCGGGGATCTCCAGCCAGGCA 187
DB 44 SerProAlaGlnAlaSerProAlaGlyThrProProGlyArgAlaSerProAlaGlnAla 63
QY 188 TCTCCAGCTGGTACCTCCAGCGCGGATCTCCAGCGCG-----GCA 232
DB 64 SerProAlaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProAlaGlnAla 83
QY 233 TCTCCAGCCAGGATCTCCAGCGCGGATCTTCGGGCTGTGGCATCTTCCAGGTCC 292
DB 84 SerProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSerLeuSerArgSer 103
QY 293 TCATCCGAGAGTCAATCCGCGGATGAGCTCGGTGACCACTCCCAACCAAGATG 352
DB 104 SerSerGlyArgSerSerSerAlaArgSerAlaSerAlaThrThrSerProThrArgVal 123
QY 353 TACCTTGTAGACCAACACAGTGGGGCTGTACCATCCGATATCTCTGCCAGGTCA 412
DB 124 TyrLeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSer 143
QY 413 GCACCAAGACCAAGGCGCACAGGAGAGGCGAGGTACGAGCCTGCGCCAGTTCACTGG 472
DB 144 AlaProAlaThrArgAlaThrArgLysSerProGlyThrSerLeuProLysPheThrTrp 163
QY 473 CGGAGGCGCAGAGACGCTTACCGCTCATCGGATCGGCTGCTCTCATTTGCCCTGATG 532
DB 164 ArgGluGlyGlnLysGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuVal 183
QY 533 GTTTCGCTCATCATCTCTTCCAGTTCTGGCAGGCGCACACAGGATCAGGTACAGAG 592
DB 184 ValSerLeuIleIleLeuPheGlnPheTrpGlnIleYHlaThrGlyIleArgTyrLysGlu 203
QY 593 CAGAGGAGAGCTGTCCACAGACGCTGTTCGCTGACGCGGAGTGTGACTGCAACTG 652
DB 204 GlnArgLysSerCysProLysHlaValAlaArgCysAepGlyValAlaAepCysLysLeu 223
QY 653 AAAGTGACAGAGCTGGGCTCGTGAAGTTGACTGGAGCAAGTCTGCTTAAATGTAC 712
DB 224 LysSerAepGluLeuGlyCysValArgPheAepTrpApySerLeuLeuIleLys 243
QY 713 TCTGAGTCTCCCATCATGATGGCTTCCATCTGTAGCAGCACTGGATGACTCTACTCA 772
DB 244 SerGlySerSerHlaGlnTrpLeuProIleCysSerSerAenTrpAenAepSerLysSer 263
QY 773 GAGAAAGCTGCCAGCAGCTGGGTTTGGAGAGTCTACCGGACCAACGAGGTGGCCAC 832
DB 264 GluLysThrCysGlnGlnLeuGlyPheGlnSerAlaHlaArgThrThrGluValAlaHla 283
QY 833 AGGATTTTGCACACAGCTTCTCAATCTGATGATACACTCCACATCCAGAAAGCTC 892
DB 284 ArgAepPheAlaAenSerPheSerIleLeuArgTyrAenSerThrIleGlnGlnSerLeu 303
QY 893 CACAGTCTGAATGCCCTTCCACGCGGTATATCTCCTCCAGTCTTCCACTGGGAGCTG 952

Db 304 HieArSerGluCyProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeu 323
QY 953 AGGGCCATGACCGGGCGGATCTGGGAGGGGCGCTGGCTCGATACAGTGGCCCTTG 1012
Db 324 ArgAlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTrpProTyr 343
QY 1013 CAAGTGAAGTGTGCACTTGGGACCAACCACATCTGTGGAGGCAAGCTCATTTGACGCCAG 1072
Db 344 GlnValSerLeuHisPheGlyThrThrHisIleGlySerIleGlyThrLeuIleAspAlaGln 363
QY 1073 TGGGTGCTCACTGCGCCGCACTGCTTCTGTGTAACCGGAGAAAGTCTTGGAGGCTGG 1132
Db 364 TrpValLeuThrAlaAlaHisCysPhePheValThrArgGlyValLeuGlnGlyTyr 383
QY 1133 AAGGTGAACCGGGCACCAGCAACCTGCACAGTGGCTTGGAGGCTTCATTTGGCCGAG 1192
Db 384 LysValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlnAlaAlaSerIleAlaGln 403
QY 1193 ATCATCATCAACAGCAATTTACACCGATGAGGAGGACGACTATGACATCGCCCTCATGCGG 1252
Db 404 IleIleIleAsnSerAsnTyrThrAspGlnGlnAspAspTyrAspIleAlaLeuMetArg 423
QY 1253 CTGTCCAGGCCCTGACCTGTGCTCCGCTCACATCCACCTGTGCTTGCCTCCCATGCAATGA 1312
Db 424 LeuSerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGly 443
QY 1313 CAGACCTTTAGCTCATATGAGACCTGCTGCATCAGAGCTTTGGCAAGACGAGGAAACA 1372
Db 444 GlnThrPheSerLeuAsnGlnThrCysTrpIleThrIleGlyLysTrpArgGlnThr 463
QY 1373 GAGGACAGACATCCCTCCCTCCGAGGAGTGCAGGTCATCTCATCGACTTCAAGAAA 1432
Db 464 AspAspLysThrSerProPheLeuArgGlnValGlnValAsnLeuIleAspPheLysLys 483
QY 1433 TGCATGACTACTTGTGTCTATGACAGTTACTTACCCCAAGATGATGTGTGTGGGAGAC 1492
Db 484 CysAsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAsp 503
QY 1493 CTTGTGGGGGCGAGACTCTGTCCAGGAGACAGCGGGGGGCTCTTGTGTGTGAGCAG 1552
Db 504 LeuArgIleGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGln 523
QY 1553 AACAAACGGCTGCTCCGCGGAGGTGACCAACGCTGGGGGCAAGGCTGTGGCAGAGAAAC 1612
Db 524 AsnAsnArgTyrTyrLeuAlaGlyValThrSerTyrGlyCysGlyGlnAsn 543
QY 1613 AAACCTGATGTATGACCAAGATGACAGAAATTCTTCCCTGATTTACAGCAAGATGAG 1672
Db 544 LysProGlyValTyrThrLysValThrGlnValLeuProTyrIleTyrSerLysMetGln 563

RESULT 10
US-10-428-275-104
; Sequence 104, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428,275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390

; PRIOR FILING DATE: 2001-02-23
; APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: CuraSequit version 0.1
; SEQ ID NO 104
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-275-104
Alignment Scores:
Pred. No.: 2,02e-171 Length: 586
Score: 2948.50 Matches: 554
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 0
Query Match: 88.5% Indels: 5
DB: 4 Gaps: 1
US-10-806-370-11 (1-1748) x US-10-428-275-104 (1-586)
QY 11 ATGAGAGGAGACACCGGGAATGCATCTCCAGCAAGAACACTTTCAGCTTGAGACATCT 70
Db 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGCACTCCAGCTGGGAGACACTCCGAGCCGGGCACTCTCCAGCCAGGCACTCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProGlyLysArgAlaSerProAlaGlnAlaSer 40
QY 131 CCAGCCAGGCACTCCAGCTGGGAGACACTCCGAGCCGGGCACTCTCCAGCCAGGCACTCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProGlyLysArgAlaSerProAlaGlnAlaSer 60
QY 191 CCAGCTGTACACCTCCAGGCGGGCACTCTCCAGGCGG-----GCATCT 235
Db 61 ProAlaGlyThrProProGlyArgAlaSerProGlyLysArgAlaSerProAlaGlnAlaSer 80
QY 236 CCAGCCAGGCACTCCAGGCGGGGCACTCCGAGCTTGCATCATCTTCCAGGCTCTCA 295
Db 81 ProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSerLeuSerAspSer 100
QY 296 TCCGGCAGATCATATCCGCGAGGTCAGCTCGGTGACCACTCCCAACAGAGTGTAC 355
Db 101 SerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
QY 356 CTTGTTAGACCAACACCAAGTGGGGGCTGTACCCATCCGATCATCTCTCCAGGTACGA 415
Db 121 LeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
QY 416 CCAGCAACCGGGCCACCGAGGAGAGCCGAGTACGACCTGCGCAAGTCCACTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgLysSerProGlyThrSerLeuProLysPheThrTyrArg 160
QY 476 GAGGGCAGAGAGAGTACCGGCTCATCGGGTGTGCTCTCCCTCATTTGCCGTGGTGT 535
Db 161 GlnGlyGlnGlySerLeuProLeuIleGlyCysValLeuLeuLeuIleAlaValVal 180
QY 536 TCGCTCATCATCTCTTCCAGTTCTGGCAGGGCCACACAGGATCAGTACCAAGAGCAG 595
Db 181 SerLeuIleLeuPheGlnPheTrpGlnGlyHisThrGlyLeaGlyTyrLysGlnGln 200
QY 596 AGGAGAGCTGTCCCAAGCACCGCTGTTCGCTGACCGGGTGTGAGCTCAAGCTGAAG 655
Db 201 ArgGlnSerCysProLysHisAlaValArgCysAspGlyValAlaAspCysLysLeuLys 220
QY 656 AGTACAGAGCTGGGCTGCGTGAAGTTGACTGGGACAAAGTCTGCTTAAATCTACTCT 715
Db 221 SerAspGlnLeuGlyCysValArgPheAspTyrAspLysSerLeuLeuLysIleTyrSer 240
QY 716 GGGTCTCTCCATCAGTGGCTTCCATCTGTAGACAGCACTGGAATGACTCTTACTGAG 775
Db 241 GlySerSerHisGlnThrLeuProIleCysSerSerAsnThrPheAspSerTyrSerGln 260

```
QY 776 AAGACCTGCCAGCAGCTGGGTTTGAGAGTCTTACCGGAGACCGAGAGTTGCCCAAG 835
Db 261 LyrthrCysgInglInleuGlyPhegluSerLanIsargThrGluValAlaHisArg 280
QY 836 GATTTCGCAACAGCTTCTCAATCTTGAGATACCACTCCACATCCAGGAAGCTCCAC 895
Db 281 AspheAlaAsnSerPheSerLileuArgLysAsnSerThrIleGlnGluSerLeuHis 300
QY 896 AGGTCTGAATCCCTTCCACAGCGATATATCTCCCTCCAGTGTCCCACTGGCACTGAG 955
Db 301 ArgSerGluCysProSerGlnArgLysThrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
QY 956 GCCATGACCGGGCGGATCTGTGGAGGGCGCTGGCTCGGATACCAAGTGCCTTGGCAA 1015
Db 321 AlameThrArgLysIleValGlyValAlaLeuAlaSerAspSerLysTrpProTrpGln 340
QY 1016 GTGAGTCGCACTTCGGGACCAACCATCTGTGAGAGGACGCTCATTTGAGCCCAAGTG 1075
Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnTrp 360
QY 1076 GTGCTCACTGGCGGCCACTGCTTCTTCGTGACCCGGGAGAAAGTCTGAGAGGCTGGAAG 1135
Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgLysValLeuGlnGlyTrpLys 380
QY 1136 GTGTACCGGGGACCCAGCAACCTGACACCACTTGGAGGAGCTCCATTTGCCGAATC 1195
Db 381 ValLysArgLysGlyThrSerAsnLeuHisIleGlnLeuProGlnAlaAlaSerLileAlaGlnIle 400
QY 1196 ATCATCAACAGCAATTTACACGATGAGGAGGAGCACTATGACATCGCCCTCATCGGCTG 1255
Db 401 IleIleAsnSerAsnLysThrAspLysGlnLysAspLysArgLysIleAlaLeuMetArgLeu 420
QY 1256 TCCAGGCCCTTGACCTGTCCGCTCACATCCACCTGCTGCTGCCATGCATGACAG 1315
Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeuProMetHisGlyGln 440
QY 1316 ACCTTTAGCTCAATGAGACCTGTGTGATCAACAGGCTTTGGCAAGAACGAGGAGACAGAT 1375
Db 441 ThrPheSerLeuAsnGlnThrCysThrLysThrArgLysPheGlyLysThrArgGluThrAsp 460
QY 1376 GACAAGACATCCCTTCCTCCGGGAGGTGAGGATCATCTCATCGACTTAAAGATGC 1435
Db 461 AspLysThrSerProPheLeuArgLysValGlnValAsnLeuIleAspPheLysLysCys 480
QY 1436 AATGACTACTGTGCTATGACAGATTACTTACCCCAAGATGATGTGCTGGGACCTT 1495
Db 481 AsnAspLysLeuValLysAspSerLysLeuThrProArgMetMetCysAlaGlyAspLeu 500
QY 1496 CTGGGGGCGAGAGACTCCTGCGCAGGAGACAGCGGGGGGCTTGTGTGTGACCAAGC 1555
Db 501 ArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGlnAsn 520
QY 1556 AACCGGTGTACTCTGGGAGAGGTGACACGCTGGGGGACAGGCTGTGGCCAGAGAAACAA 1615
Db 521 AsnArgLysTrpLysLeuAlaGlyValThrSerTrpGlyThrArgLysCysGlyGlnArgAsnLys 540
QY 1616 CTTGGTGTGACCAAGATGACAGAAATTCTTCCCTGATTTACAGAGATGAGAG 1672
Db 541 ProGlyValLysTrpThrLysValThrGluValLeuProTrpLysLysSerLysMetGln 559
RESULT 11
US-10-428-275-124
; Sequence 124, Application US//10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US//10/428, 275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
```

```
QY 11 ATGAGAGGAGACAGCCAGGGAATGCATCTCCAGCAAGAACACCTTGAGTGCAGATCT 70
Db 1 MetGluArgAspSerHisIleGlyAsnAlaSerProLalaArgThrProSerAlaGlyAlaSer 20
QY 71 CCAGCCAGGCAATCTCCAGCTGGGACACCTCCAGGCGGGGCAATCTCCAGCCCAAGCATCT 130
Db 21 ProLalaGlnAlaSerProLalaGlyThrProProGlyArgAlaSerProLalaGlnAlaSer 40
QY 131 CCAGCCAGGCAATCTCCAGCTGGGACACACCTCCGGGCGGGGCAATCTCCAGCCCAAGCATCT 190
Db 41 ProLalaGlnAlaSerProLalaGlyThrProProGlyArgAlaSerProLalaGlnAlaSer 60
QY 191 CCAAGCTGTACACCTCCAGGCGGGGCAATCTCCAGGCGGG-----GCATCT 235
Db 61 ProLalaGlyThrProProGlyArgAlaSerProGlyArgAlaSerProLalaGlnAlaSer 80
QY 236 CCAGCCAGGCAATCTCCAGGCGGGGCAATCTCCGGCTTGGCATCACTTCCAGGCTCTCA 295
Db 81 ProLalaGlnAlaSerProLalaArgAlaSerProLalaLeuAlaSerLeuSerArgSerSer 100
QY 296 TCCGGGAGGTCATCATCGGCGAGGTCAGCTCGGTGACAACTCCCAACCAAGAGTGTAC 355
Db 101 SerLysArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValLys 120
QY 356 CTTGTTAGACCAACCAAGTGGGGGCTGTACCCATCCGATCATCTCTGCGCAGGTGACCA 415
Db 121 LeuValArgLalaThrProValGlyAlaValProLalaArgSerSerProLalaArgSerAla 140
QY 416 CCAGCAACCAAGGCGCACCAAGGAGAGCCAGGTCAGAGCTGCCCAAGTTCACTGGCGG 475
Db 141 ProLalaThrArgAlaThrArgLysSerProGlyThrSerLeuProLysPheThrTrpArg 160
QY 476 GAGGGCCAGAGAGCAGTACCGCTCATGGGGGCGTCTCTCCCTCATTTGCGTGGTGGT 535
Db 161 GlnGlyGlnLysGlnLeuProLeuIleGlyCysValLeuLeuLeuIleAlaLeuValVal 180
QY 536 TCGCTCATCATCTCTTCCAGTTCGTGGCAGGCGCACAGAGATCAAGTCAAGAGCAG 595
Db 181 SerLeuIleIleLeuPheGlnPheTrpGlnGlyHisThrGlyLysArgLysLysGlnGln 200
QY 596 AGGAGAGCTGTCCCAAGCACGCTGTCTGTGACCGGGGTGTGAGCTCAAGCTGAAG 655
```



```

Db      201 ArgGluSerCysProIyShiAlaValArgCysAspIyValValAspCysIyLeuLys 220
QY      656 AGTGAAGAGTGGCTGGCTGGAGTTGACTGGAGCAAGCTCTGGCTTAATATCTCT 715
        |||
Db      221 SerAspIyLeuGlyCysValArgPheAspIyPheAspIySerLeuLysIleTyrSer 240
QY      716 GGGTCCCGCATCAGTGGCTTCCATCTGTAGAGCAGCAATGGAAATGACTCTACTGAG 775
        |||
Db      241 GlySerSerIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 260
QY      776 AAGACCTGACAGCAGTGGCTTGCAGAGTCTCAGCAGCAACCGAGTGGTCCAGAG 835
        |||
Db      261 LysThrCysGlnGlnIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 280
QY      836 GATTTTCCCAACAGCTTCTCATATTGAGATACACTCCACATCCAGAAAGCTCCAC 895
        |||
Db      281 AspPheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIleGlnIshIshIsh 300
QY      896 AGGTCTGAATGGCTTCCAGCGGTATATCTCCCTCAGATGTTCCCATGCGGAGT 955
        |||
Db      301 ArgSerIyCysProSerGlnArgTyrIleSerLeuGlnCysSerIshCysGlyLeuArg 320
QY      956 GCCATGACCGGCGGATCTGTGGAGGCGCTGGCTCGGATAGCAAGTGGCTTGCAA 1015
        |||
Db      321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerIyStrProIyPheGln 340
QY      1016 GTGAGTGTGCACTTGGGACCAACCATCTGTGGAGGCAAGCTCATTTAGCCCACTG 1075
        |||
Db      341 ValSerLeuShIshPheGlyThrIshIshIleCysGlyIshIshIshIshIshIshIsh 360
QY      1076 GTGCTGACGCGCGGCACTGCTTCTGTGACCGGAGAAAGCTCGAGGAGCTGAG 1135
        |||
Db      361 ValLeuThrAlaAlaIshCysPhePheValThrArgLeuValLeuGlnGlyIshIsh 380
QY      1136 GTGTACCGGCGGCAACCAACCATCTGTGGAGGCAAGCTCATTTAGCCCACTG 1195
        |||
Db      381 ValTyrAlaGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 400
QY      1196 ATGATCAACAGCAATTAACCCGATGAGAGAGAGCATATGACATGCCCTCAGGCTG 1255
        |||
Db      401 IleIleAsnSerAsnTyrThrAspGlnGluAspAspIleAlaIshIshIshIshIsh 420
QY      1256 TCGAAGCCCTGACCCCTGCTGCTGACATCAACCCCTGCTGCTGCTGCTGCTGCTG 1315
        |||
Db      421 SerIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 440
QY      1316 ACCTTAGCTCAATGAGACCTGCTGATGACAGGCTTTGGCAAGACGAGGAGACAGAT 1375
        |||
Db      441 ThrPheSerLeuAsnGlnIshIshIshIshIshIshIshIshIshIshIshIshIsh 460
QY      1376 GACAAAGACATCCCTCTCTCCGAGAGTGCAGGTCATCTCATGCACTTCAAGAAATGC 1435
        |||
Db      461 AspIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 480
QY      1436 AATGACACTTGTGCTTGAAGATTAACCTTAACCCCAAGATGATGTTGCTGGGAGACTT 1495
        |||
Db      481 AsnAspIyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 500
QY      1496 CGTGGGAGGAGAGACTCTGACAGGAGAGAGAGCGGCGGCTTGTGCTGAGCAAGAC 1555
        |||
Db      501 ArgGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 520
QY      1556 AACCGCTGTAACCTGAGGCTGTCAACAGCTGGGAGCAAGGCTGTGGCCAGAGAAACAA 1615
        |||
Db      521 AsnArgIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 540
QY      1616 CTGGTGTGACCAACCAAGGAGAGAGATCTTCCCTGAGATTTACAGACAGATGAG 1672
        |||
Db      541 ProGlyValIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 558

```

RESULT 12
US-10-428-275-142
; Sequence 142, Application US/10428275

```

; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvaraz et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428,275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: CuroSeqIsh version 0.1
; SEQ ID NO 142
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-275-142

Alignment Scores:
Pred. No.: 4.07e-171 Length: 586
Score: 2943.50 Matches: 553
Percent Similarity: 98.9% Conservative: 0
Best Local Similarity: 98.9% Mismatches: 1
Query Match: 88.3% Indels: 5
DB: 4 Gaps: 1

US-10-806-370-11 (1-1748) x US-10-428-275-142 (1-586)
QY      11 ATGAGAGGAGGAGCCAGGGAATGATCTCCAGCAAGAACCTTCAGCTGAGCATCT 70
        |||
Db      1 MetGlnArgAspSerIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 20
QY      71 CCAGCCGAGCATCTCCAGCTGGGACACTCCAGCCGAGCATCTCCAGCCGAGCATCT 130
        |||
Db      21 ProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 40
QY      131 CCAGCCGAGCATCTCCAGCTGGGACACTCCAGCCGAGCATCTCCAGCCGAGCATCT 190
        |||
Db      41 ProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 60
QY      191 CCAGCTGTACCTCCAGGCGGAGCATCTCCAGGCGG-----GCATCT 235
        |||
Db      61 ProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 80
QY      236 CCAGCCGAGCATCTCCAGGCGGAGCATCTCCAGCTGTGACATCTTCAGGCTCTCA 295
        |||
Db      81 ProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 100
QY      296 TCCGCGAGGATCATCTCCAGGCTGAGCACTCCGCTGCAACCTCCCAACCAAGGTGAC 355
        |||
Db      101 SerGlyArgSerSerSerIshIshIshIshIshIshIshIshIshIshIshIshIsh 120
QY      356 CTGTTAGAGCAACACAGTGGGCGGTGTACCATCCGATCATCTCTGCGAGGTACGA 415
        |||
Db      121 LeuValArgAlaIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 140
QY      416 CCAGCAACAGGCGCACAGGAGAGAGCCAGGTACGAGCTGCCCAAGTTCACTGCGCG 475
        |||
Db      141 ProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 160

```

QY 476 GAGGGCCAGAGACAGCTACCGCTCATCGGGTCGGTCTCTCTCATTTGCCCTGGTGGTT 535
DB 161 G1UG1YGIInlysgInleuProleu1IeGIYcyValleuLeu1IeAla1eVal1 180
QY 536 TCCGTATCATTCCTTCCTCCAGTTCTGGCAGGGCCACACAGGAGTACAGTACAGAGACAG 595
DB 181 Serleu1Ie1leuPheGInPhe1rPGLnGIYH1eThrGIYleAryTYrLysGIIn 200
QY 596 AGGAGAGACTGTCCCAAGACAGCTGTTCGCTGTGACCGGGGTGGTGACTGCAAGCTGAAG 655
DB 201 ArgG1uSerCyseProlyeH1sAlaValArgCyseAspGIYValAlaAspCyLysLeuLys 220
QY 656 AGTACAGAGCTGGGCTGGTGGAGGTTGACTGGACAAAGTCTGTCTTAAATCTACTCT 715
DB 221 SerAspGIYleuGIYCyValArgPheAsp1rAspLysSerLeuLys1IeTYrSer 240
QY 716 GGGGCTCCCATCATGTGGTTCCTCATGTATGACAGCAATCTGGATGACTCTACTCATAGAG 775
DB 241 G1YSerSerH1sGIIn1rPleuPro1IeCyseSerSerAsn1rAspAspSerTYrSerGIIn 260
QY 776 AAGACTGACAGACAGCTGGGTTTGAAGATGCTCACCGGCAACACGAGTTGCCACAGG 835
DB 261 LysThrysgInleuGIYPhelG1uSerAlaH1eArgH1rThrGIYValAlaH1eArg 280
QY 836 GATTTCACCAACAGCTTCTCAATTTTGAATATCAACTCCACCATCCAGGAAAGCTTCAC 895
DB 281 AspPheAlaAsnSerPheSer1IeLeuArgTYrAsnSerThr1IeGIInG1uSerLeuH1s 300
QY 896 AGGCTGGAATGGCTTCCCAAGCGGTATATCTCCCTCCAGTGTTCCTCCATGGGACTGAG 955
DB 301 ArgSerGIYCyseProSerGIYArgTYrLaseSerLeuGIYCyseSerH1sCyseGIYLeuArg 320
QY 956 GCGATGACCGGGCGAGATCGTGGAGAGGGCGCTGGCTCGGATAGCAAGTGGCTTGGCAA 1015
DB 321 AlaMetHrGIYArg1IeValGIYAla1eLeuAlaSerAspSer1rTYrPro1rPGLn 340
QY 1016 GTGAGTGTGCACTTGGGACACCCACATCTGTGAGAGGACGCTCATTTAGCGCCAGTGG 1075
DB 341 ValSerLeuH1sPheGIYThrThrH1s1IeCyseGIYGIYThrLeu1IeAspAlaGIIn1rP 360
QY 1076 GTGCTACATCGCGGCACTGCTTCTGTGACCGGGAGAAAGTCTGGAGGGCTGGAG 1135
DB 361 ValLeuThrAlaAlaH1sCysePhePheValThrArgGIYVal1IeGIYGIYTYrLys 380
QY 1136 GTGATACCGGGACACAGCAACTGACACAGTTGCTGAGGAGCGCTCATTTGCCAGATC 1195
DB 381 ValTYrAlaGIYThrSerAlaLeuH1eGIInleuProGIYAlaAlaSer1IeAlaGIIn1Ie 400
QY 1196 ATCATCAACAGCAATTAACCCGATGAGAGAGACATATGACATGCGCTCATGCGGCTG 1255
DB 401 Thr1IeAsnSerAsnTYrThrAspGIYGIYAspAspTYrAsp1IeAla1eMetArgLeu 420
QY 1256 TCCAGAGCCGTGACCGCTCGGCTCATATCCAGCTGTGGCTTCCCAAGTACAGAGCAG 1315
DB 421 SerTyseProleuThrLeuSerAlaH1s1IeH1sPro1IeCyseLeuProMetH1sGIYGIIn 440
QY 1316 ACCTTTACCTCATATGAGACTGCTGATACAGAGCTTGTGCAAGACAGGAGAGACAGAT 1375
DB 441 ThrPheSerLeuAsnGIYThrCyse1rP1IeThrGIYPhelGIYTYrHrArgGIYThrAsp 460
QY 1376 GACAAAGCATCCCTCTCTCCGGAGAGTGCAGTCAATCTCATGCACTTCAAGAAATGC 1435
DB 461 AspTyseThrSerProPheLeuArgGIYValGIYVal1AsnLeu1IeAspPheLysLysCyse 480
QY 1436 AATGACACTTGTGCTATGACAGTTACCTTACCCCAAGATGATGTATGCTGGGAGACTT 1495
DB 481 AsnAspTyseLeuValTYrAspSerTYrLeuThrProArgMetCysValGIYAspLeu 500
QY 1496 CGTGGGGGAGAGACTCTGACAGGAGAGACAGCGGGGGCTCTTGTGCTGACAGAC 1555
DB 501 ArgGIYGIYArgAspSerCyseGIYGIYAspSerGIYGIYProleuValCyseGIYGIYAsn 520
QY 1556 AACCGTGTGACTGTGACAGGTGTGCAACAGCTGGGGCACAGGCTGTGGCCAGAGAAACAA 1615

DB 521 AsnArgTYrTYrLeuAlaGIYValThrSerTYrGIYThrGIYCyseGIYGIYArgAsnLys 540
QY 1616 CCGGTGTGTACACCAAGTACAGAGAGTCTTCCCTGGATTTACAGACAGATGAG 1672
DB 541 ProGIYValTYrThrLysValThrGIYValleuPro1rP1IeTYrSerLysMetGIIn 559

RESULT 13
US-10-428-275-140
Sequence 140, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428,275
PRIOR APPLICATION NUMBER: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseqdist version 0.1
SEQ ID NO 140
LENGTH: 586
TYPE: PRT
ORGANISM: Homo sapiens
US-10-428-275-140

Alignment Scores:
Pred. No.: 4,686-171 Length: 586
Score: 2942.50 Matches: 553
Percent Similarity: 98.9% Conservative: 0
Best Local Similarity: 98.9% Mismatches: 1
Query Match: 88.3% Indels: 5
DB: 4 Gaps: 1

US-10-806-370-11 (1-1748) x US-10-428-275-140 (1-586)

QY 11 ATGGAGAGGACAGACGAGGAATGCATCTCCAGCAAGAACACCTTCAGCTGAGCATCT 70
DB 1 MetGIYArgAspSerH1sGIYAsnAlaSerPro1IaArgH1rProSerAlaGIYAlaSer 20
QY 71 CCAAGCCAGGCAATCTCCAGCTGGAGCACTTCCAGCCGGGCAATCTCCAGCCAGCATCT 130
DB 21 Pro1IaGIYAlaSerPro1IaGIYThrProProGIYArg1IaSerPro1IaGIYAlaSer 40
QY 131 CCAAGCCAGGCAATCTCCAGCTGGAGCACTTCCAGCCGGGCAATCTCCAGCCAGCATCT 190
DB 41 Pro1IaGIYAlaSerPro1IaGIYThrProProGIYArg1IaSerPro1IaGIYAlaSer 60
QY 191 CCAAGCTGTACACTCCAGGCGGGGCAATCTCCAGGCGG-----GCATCT 235
DB 61 Pro1IaGIYThrProProGIYArg1IaSerProGIYArg1IaSerPro1IaGIYAlaSer 80
QY 236 CCAAGCCAGGCAATCTCCAGCCGGGCAATCTCCAGCTTGGCATCATTTCCAGGCTCTCA 295
DB 81 Pro1IaGIYAlaSerPro1IaArg1IaSerPro1IaLeuAlaSerLeuSerArgSerSer 100
QY 296 TCCGGCAGGTCATCATCTCCGCAAGGTACAGCTCGGTACCAACTCCCAACCAAGGTATAC 355

Db 101 SerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSerProThrArgValTyr 120
Qy 356 CTGTTGAGAGAAACACAGTGGGGCTGTACCCATCCGATATCTCTGCGAGGTGACA 415
Db 121 LeuValArgAlaThrProValGlyAlaValProIleArgSerSerProAlaArgSerAla 140
Qy 416 CCAAGCAACAGGGGACACAGGAGAGAGCCAGGTACAGACCTGCGCCAGTTCACTGGCGG 475
Db 141 ProAlaThrArgAlaThrArgGlySerProGlyThrSerLeuProGlySerPheThrTyrArg 160
Qy 476 GAGGCGCAAGAGACGATACCGCTCATCGGGTGGCTGCTCTCTCATTTGCCCTGGTGGT 535
Db 161 GlyGlyGlnGlySerGlnLeuProLeuIleGlyValLeuLeuLeuIleAlaLeuValVal 180
Qy 536 TCCGTTATCATCTCTCTTCCAGTTCTGGCAGAGGCCACACAGGATCAGGTACAAAGAGCAG 595
Db 181 SerLeuIleIleLeuPheGlnPheThrGlnGlyIleThrGlyIleArgGlyValGlnGln 200
Qy 596 AGGAGAGACTGTCCCAAGACGCTGTTGCTGTGACGGGGTGGTGGACTGCAACTGAAAG 655
Db 201 ArgGlySerCysProGlySerGlnAlaValArgCysArgProGlyValValArgCysLeuLeu 220
Qy 656 AGTGCAGAGCTGGGCTGCGTGGAGGTTTACTGGAGCAAGTCTGCTTAAATCTACTCT 715
Db 221 SerArgGlyLeuGlyCysValArgPheArgPheArgPheArgPheArgPheArgPheArg 240
Qy 716 GGGTCTCCATCATGAGTGGCTTCCATCTGTAGCAGCACTGGAAATGACTCTTACTCAGAG 775
Db 241 GlySerSerHisGlnTyrLeuProIleCysSerSerHisTyrPheArgPheArgPheArg 260
Qy 776 AAGACCTGCGACAGCTGGGTTTGGAGAGTCTCACCGGAAACCGAGGTTGCCACAGAG 835
Db 261 LysThrCysProGlnLeuGlyPheGlnSerAlaHisArgGlnThrGlnValAlaHisArg 280
Qy 836 GATTTCGCAACAGCTTCTCATCTTGAATACATCCACATCCAGCAAGCAAGGCTCCAC 895
Db 281 AspPheAlaHisSerPheSerIleLeuArgGlyThrAsnSerThrIleGlnGlnSerLeuHis 300
Qy 896 AGGTCTGAATGCCCTTCCAGCGGATATCTCCCTCCAGTGTCTCCACTGCGGACTGAAG 955
Db 301 ArgSerGlyCysProSerGlnArgTyrIleSerLeuGlnCysSerHisCysGlyLeuArg 320
Qy 956 GCCATGACCGGGCGGATCGTGGGAGGGGCGCTGGCTCGATACAGTGGCTTGGCAA 1015
Db 321 AlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSerLysTyrProIleProGln 340
Qy 1016 GTGAGTGTGCACTTGGCAGCACCACATCTGTGGAGGAGGCTCATTTGACGCCAGTGG 1075
Db 341 ValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeuIleAspAlaGlnThr 360
Qy 1076 GTGCTCACTGCGCGGCACTGCTTTTGTGTGACCCGGGAAAGGTTCTGGAGGCTGAAAG 1135
Db 361 ValLeuThrAlaAlaHisCysPhePheValThrArgGlyValLeuGlyGlyTyrLys 380
Qy 1136 GTGAGCGGGGACACAGCACTGSCACAGTGGCTGAGGAGGCTCATTTGCGGAGATC 1195
Db 381 ValTyrAlaGlyThrSerAsnLeuHisGlnLeuProGlyAlaAlaSerIleAlaGlnIle 400
Qy 1196 ATCATCAACAGCAATTCACGATGAGAGAGCACTATGACATCGCCTCATCGGCTG 1255
Db 401 IleLeuHisSerAsnTyrThrArgGlyGlyAspArgLysIleAlaLeuMetArgLeu 420
Qy 1256 TCCAGGCGCTGACCTGTGCGCTCACATCCACCTGCTTGGCTCCCGCATGAGCAAG 1315
Db 421 SerLysProLeuThrLeuSerAlaHisIleHisProLacLysLeuProMetHisGlyGln 440
Qy 1316 ACCCTTGAGCTGATGAGACCTGCTGATACAGAGCTTGGCAGAGCAGGAGAGCAAGAT 1375
Db 441 ThrPheSerLeuAsnGlnThrCysTyrIleThrGlyPheGlyLysThrArgGlnThrAsp 460
Qy 1376 GACAAGACATCCCTCTCTCCGAGAGGTGAGGTCAATCTCATCGAATTCAGAAATGAC 1435
Db 461 AspArgThrSerProPheLeuArgGlyValGlnValAsnLeuIleAspPheLysLysCys 480

Qy 1436 AATGACTACTTGGCTATGACAGTTACTTACCCCAAGAGATGATGTGGGGACCTT 1495
Db 481 AsnAspTyrLeuValTyrAspSerTyrLeuThrProArgMetCysAlaGlyAspLeu 500
Qy 1496 GTGGGGGACAGAGACTCTGTCAGAGGAGACAGCGGGGGCTCTTGTCTGTGACAGAAC 1555
Db 501 ArgGlyGlyThrArgAspSerCysGlnGlyAspSerIleGlyProLeuValCysGlyGlnAsn 520
Qy 1556 AACCGCTGTACTCTGCGAGGTGTACACAGTGGGGCACAGGCTGTGGCCAGAGAAACAA 1615
Db 521 AsnArgTyrTyrLeuAlaGlyValThrSerProGlyThrGlyCysGlyGlnArgAsnLys 540
Qy 1616 CTTGGTGTGACACCAAGTGACAGAGTCTTCCGATTTACAGACAGATGAG 1672
Db 541 ProGlyValTyrThrLysValThrGlnValLeuProIleTyrSerLysMetGln 559

RESULT 14
US-09-888-615-104
Sequence 104, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOMMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEDEBEL, SEAN
APPLICANT: CHARNOJCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888, 615
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 104
LENGTH: 537
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-104

Alignment Scores:
Pred. No.: 7,07e-161 Length: 537
Score: 2775.50 Matches: 526
Percent Similarity: 92.0% Conservative: 0
Best Local Similarity: 92.0% Mismatches: 1
Query Match: 83.3% Indels: 45
DB: 3 Gaps: 2

US-10-806-370-11 (1-1748) x US-09-888-615-104 (1-537)

Qy 11 ATGAGAGAGGACAGCAGGGAATGATCTTCAGACAAACAACCTTGAGTGGAGCATCT 70
Db 1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
Qy 71 CCAAGCCAGGATCTCAGCTGGGACACTTCAGAGCGGGGATCTTCCAGCCAGGATCT 130
Db 21 ProAlaGlnAlaSerProAlaGlyThrProProGlyThrArgAlaSerProAlaGlnAlaSer 40
Qy 131 CCAAGCCAGGATCTCAGCTGGGACACTTCAGAGCGGGGATCTTCCAGCCAGGATCT 190
Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyThrArgAlaSerProAlaGlnAlaSer 60
Qy 191 CCAAGCTGTACACTCCAGGCGGGGATCTTCAGAGCGG----- 229
Db 61 ProAlaGlyThrProProGlyThrArgAlaSerProGlyThrArgAlaSerProAlaGlnAlaSer 80
Qy 230 -----GCATCTCAGCGCCAGGCACTTCAGAGCGGGGATCTCCGGCTTGGCATCA 280
Db 81 ProAlaGlnAlaSerProAlaGlnAlaSerProAlaArgAlaSerProAlaLeuAlaSer 100
Qy 281 CTTTCCAGTCTCATCTCCGAGAGGTATCATCCGCCAGCTCAGCTGCTGGTGAACAACCTCC 340

```

Db      |||
101 LeuSerArgSerSerSerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSer 120
Qy      |||
341 CCAACGAGAGTATCTCTTTAGAGCAACACGATGGGGGCTGTAACCATCCGATCTATCT 400
Db      |||
121 ProThrArgValIleValIleValAlaThrProValGlyAlaValProIleArgSerSer 140
Qy      |||
401 CTGCGCAGGTGAGCAGCAGCAACGAGGCGCACGAGGAGAGCCGAGGTACGAGCTGCC 460
Db      |||
141 ProIleArgSerAlaProAlaThrArgAlaThrArgIleSerPro----- 155
Qy      |||
461 AAGTTCACTGGCGGAGGCGCAGAGACGCTACCGCTCATGGGTGGCTCTCTCTC 520
Db      |||
155 ----- 155
Qy      |||
521 ATTGCCCTGGGTGTTTGGCTCATCTCTTCAGATTCTGGCAGGGGCAACAGGGATC 580
Db      |||
156 -----ValGlnPheTrpGlnGlyHisThrGlyIle 165
Qy      |||
581 AGGTACAAGGAGCAGAGGAGAGAGCTGTCCAGACGCTGTTCCGCTGTGACGGGGTGGT 640
Db      |||
166 ArgTyrIleGlnGlnArgIleSerCysProIleHisAlaValArgCysAspGlyValVal 185
Qy      |||
641 GACTGCAAGCTGAAGATGACAGCTGGGCTGGCTGGATTGACTGGAGACAATCTCTG 700
Db      |||
186 AspCysIleLeuIleSerArgGlnLeuGlyCysValAlaArgPheAspTrpAspIleLeu 205
Qy      |||
701 CTTAAATCTACTTTGGATCTCTCCATCAGAGGCTCCCATCTGTAGCAGACATGGAA 760
Db      |||
206 LeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 225
Qy      |||
761 GACTCTACTCAGAGAAAGCTGCGCAGACGCTGGGTTTGAAGTGTCTACCGGACAAC 820
Db      |||
226 AspSerTyrSerGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 245
Qy      |||
821 GAGGTTCCTCCACAGGATTTTGGCACAAGCTTCTCATTTTGAATATCACTCAACATC 880
Db      |||
246 GluValAlaHisArgAspPheAlaAsnSerPheSerIleLeuArgTyrAsnSerThrIle 265
Qy      |||
881 CAGGAAGGCTCCACAGATCTGAATGCGCTTCCAGCGGATATATCTCCCTCAGATGTC 940
Db      |||
266 GlnIleSerLeuHisArgSerGlnCysProSerGlnArgTyrIleSerLeuGlnCysSer 285
Qy      |||
941 CACTCGGAGCTGAGGAGCCATGACCGGCGAGATCTGGGAGGAGGCGCTCGCATAGC 1000
Db      |||
286 HisGlySerIleuAlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSer 305
Qy      |||
1001 AAGTGGCTTGGCAAGTGAAGTCTGCACTTGGGACACCCACATCTGTGAGGACGCTC 1060
Db      |||
306 LysTrpProTrpGlnValSerLeuHisPheGlyThrThrHisIleCysGlyIleThrLeu 325
Qy      |||
1061 ATTGACGCGCAGTGGGTGCTCAGTGGCGGCCCATGCTTCTTGGTACCGGAGAAAGTCT 1120
Db      |||
326 IleAspIleGlnTrpValIleThrAlaAlaHisIleCysPhePheValIleThrGlyValAla 345
Qy      |||
1121 CTGAGAGGCTGGAAGGTGTACGCGGAGCAGCAGCACTGCAACAGTTGCTTGAAGGACCC 1180
Db      |||
346 LeuGlnGlyTrpIleValIleValIleGlyThrSerLeuHisGlnLeuProGlnAlaAla 365
Qy      |||
1181 TTCATTCCGAGATCATCATCAACAGCAATTAACCGCATGAGGAGAGCAGTATGACATC 1240
Db      |||
366 SerIleIleGlnIleIleIleIleAsnSerAsnIleThrAspGlnIleAspAspIle 385
Qy      |||
1241 GCCCTCATGGGAGCTGTCCAGACCCCTGACCTGCGCTGACATCCACCCGCTGTGCTC 1300
Db      |||
386 AlaLeuMetCysIleSerIleProLeuHisIleSerIleHisIleIleProAlaCysLeu 405
Qy      |||
1301 CCATGATGAGACAGACCTTAAAGCTCATGAGACCTGTGATCAGAGGCTTGTGCAAG 1360
Db      |||
406 ProMetHisGlyGlnThrIlePheSerIleuAlaGlnIleThrCysTrpIleThrGlyPheGlyLys 425
Qy      |||
1361 ACCAGGAGACAGATGACAGACATCCCTCTCTCGGAGAGGTGACAGTCAATCTCATC 1420

```

```

Db      |||
426 ThrArgGlnThrAspAspIleThrSerProPheLeuArgGlnValGlnValAsnLeuIle 445
Qy      |||
1421 GACTTCAAGAAATGCAATGACTACTTGATCTATGACAGTTACTTACCCCAAGGATGATG 1480
Db      |||
446 AspPheLeuIleCysAsnAspTrpLeuValIleThrAspSerTyrLeuThrProArgMetMet 465
Qy      |||
1481 TTGCTCGGGACCTTGTGGGGGAGAGACTCTCTCCAGGAGAGCAGCGGGGCTCTTT 1540
Db      |||
466 CysAlaGlyAspLeuArgGlyIleArgAspSerCysGlnGlyAspSerGlyIleProLeu 485
Qy      |||
1541 GTCTGTGACAGAACCAACCCCTGTTACTTGACAGTGTACCCAGCTGGGCGACAGGCTGT 1600
Db      |||
486 ValCysGlnGlnHisAsnAsnArgTrpTrpLeuAlaGlyAlaIleThrSerTrpGlyThrCys 505
Qy      |||
1601 GGCGCAGAAACAAACCTGTGTGTATACACCAGTACAGAAAGTTCTTCCCTGGATTATAC 1660
Db      |||
506 GlyIleHisArgAsnIleProGlyValIleThrIleValIleThrGlnValIleuProTrpIleTyr 525
Qy      |||
1661 AGCAAGATGAGAGAGCGAGGTGCGATTGAGAAATCC 1696
Db      |||
526 SerIleMetGlnSerGlnValArgPheArgIleSer 537

RESULT 15
US-10-428-275-128
: Sequence 128, Application US/10428275
: Publication No. US20040067505A1
: GENERAL INFORMATION:
: APPLICANT: Alvarez et al.
: TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
: FILE REFERENCE: 21402-585
: CURRENT APPLICATION NUMBER: US/10/428,275
: PRIOR FILING DATE: 2003-05-01
: PRIOR APPLICATION NUMBER: 09/96545
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/544511
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: 60/128514
: PRIOR FILING DATE: 1999-04-09
: PRIOR APPLICATION NUMBER: 09/569269
: PRIOR FILING DATE: 2000-05-11
: PRIOR APPLICATION NUMBER: 60/134315
: PRIOR FILING DATE: 1999-05-14
: PRIOR APPLICATION NUMBER: 09/619252
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185548
: PRIOR FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 450
: SOFTWARE: Curseseq1st version 0.1
: SEQ ID NO 128
: LENGTH: 537
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-428-275-128

Alignment Scores:
Pred. No.: 7,07e-161 Length: 537
Score: 2775,50 Matches: 526
Percent Similarity: 92,0% Conservative: 0
Best Local Similarity: 92,0% Mismatch: 1
Query Match: 83,3% Indels: 45
DB: 4 Gaps: 2

US-10-806-370-11 (1-1748) x US-10-428-275-128 (1-537)
Qy      |||
11 ATGAGAGGAGCAGACCGGAAATGATCTTCACAGAAAGAACCTTGAAGTGAAGCATCT 70
Db      |||
1 MetGlnArgAspSerHisGlyAsnAlaSerProAlaArgThrProSerAlaGlyAlaSer 20
Qy      |||
71 CCAGCCAGGACATCTCAGCTGGGAGACCTCCAGCGGGGACATCTCAGCCGACAGCATCT 130
Db      |||
21 ProAlaGlnAlaSerProAlaGlyThrProGlyIleArgAlaSerProAlaGlnAlaSer 40

```

QY 131 CCAAGCCAGGATCTCCAGCTGGAGACCTCCGGGCGGCGATCTCCAGCCAGGATCT 190
 Db 41 ProAlaGlnAlaSerProAlaGlyThrProProGlyThrAlaSerProAlaGlnAlaSer 60
 QY 191 CCAAGCTGGTACCTCCAGGCGGGCATCTCCAGCCCGG----- 229
 Db 61 ProAlaGlyThrProProGlyThrAlaSerProGlyThrAlaSerProAlaGlnAlaSer 80
 QY 230 -----GCATCTCCAGCCAGGATCTCCAGCCCGGCGATCTCCGGCTCTGGCATCA 280
 Db 81 ProAlaGlnAlaSerProAlaGlnAlaSerProAlaGlnAlaSerProAlaGlnAlaSer 100
 QY 281 CTTTCAGGCTCCATCCGGAGGCTCATCCGGCGGCTCAGGCTCGGTGACAACTCC 340
 Db 101 LeuSerArgSerSerSerGlyArgSerSerSerAlaArgSerAlaSerValThrThrSer 120
 QY 341 CCAACCAAGAGTGTACTTGTAGACAAACCAAGTGGGGCTGTACCATCCATCCATCATCT 400
 Db 121 ProThrArgValThrLeuValArgAlaThrProValGlyAlaValProIleArgSerSer 140
 QY 401 CTGGCCAGGTACAGCAACAGGAGGCGCAACGAGGAGCCAGGTACGAGCTGCC 460
 Db 141 ProAlaArgSerAlaProAlaThrArgAlaThrArgGlySerPro----- 155
 QY 461 AAGTTACCTGGGGGAGGGCCAGAACAGCTACCGCTCATCGGGTGCCTCTCTCTC 520
 Db 155 ----- 155
 QY 521 ATTGCCCTGGTGTTCGCTCATCATCTCTCCAGTTCTGGCAGGGCCACACAGGATC 580
 Db 156 -----ValGlnThrProGlnGlyThrGlyIle 165
 QY 581 AGGTACAGAGAGCAGAGGAGAGCTGTCCAGACCGCTGTCCCTGTGACGGGCTGTG 640
 Db 166 ArgGlyArgGlnGlnArgGlySerCysProIleAlaValArgCysAspGlyValVal 185
 QY 641 GACGCAAGCTGAGAGTACGAGCTGCGGCGCTGAGGTTGACTGGGAAATCTCTG 700
 Db 186 AspCysLeuLeuLeuSerSerSerGlnLeuGlyCysValArgPheAspTrpAspLeuSerLeu 205
 QY 701 CTTAAATCTACTTGGGCTCTCCCATCAGTGGCTTCCCATCTGTAGCAGCACTGGAAT 760
 Db 206 LeuValIleThrSerGlySerSerHisGlnThrLeuProIleCysSerSerAsnTrpAsn 225
 QY 761 GACTCTACTCAGAGAAAGCTGACAGAGCTGGGCTTTCGAGAGTCTCACCGGACACC 820
 Db 226 AspSerThrSerGlnThrCysGlnGlnLeuGlyPheGlnSerAlaHisArgThrThr 245
 QY 821 GAGGTGGCCACAGAGATTGGCCACAGCTTTCATCTTGAGATACAACTCCACCATC 880
 Db 246 GluValAlaHisArgAspPheAlaAsnSerPheSerIleLeuArgTyArgSerThrIle 265
 QY 881 CAGGAAAGCCTCCACAGGTCTGAATGCGCTTCCAGGGGTATATCTCCCTCAGTGTTC 940
 Db 266 GlnGlnSerLeuHisArgSerGlnCysProSerGlnArgTyIleSerLeuGlnCysSer 285
 QY 941 CACTGCGAGTACAGGCGCATAGCCGGGCGAGTCGTGGAGGGGCGCTGGCATAGC 1000
 Db 286 HisCysGlyLeuValArgAlaMetThrGlyArgIleValGlyAlaLeuAlaSerAspSer 305
 QY 1001 AAGTGGCTTGGCAGAGTGTGCACTTGGGCAACCCCACTGTGTGAGGAGCACTGC 1060
 Db 306 TyrTrpProTrpGlnValSerLeuHisPheGlyThrThrHisIleCysGlyGlyThrLeu 325
 QY 1061 ATTGACGCCAGTGGGTCTCACTGCGGCCCATCTGCTTCTGTGACCCGGGAGAGGATC 1120
 Db 326 IleAspAlaGlnTrpValLeuThrAlaAlaHisCysPhePheValThrArgGlyValVal 345
 QY 1121 CTGAGAGGCTGGAAAGGTGTACCGGGGCAACGAACTGCAACGTTGCTGAGGAGACC 1180
 Db 346 LeuGlnGlyTrpValValValValGlyThrSerAsnLeuHisGlnLeuProGluAlaAla 365

QY 1181 TCCATTGCCAGATCATCATCAACAGCAATTACACCGATGAGAGAGCACTATAGATC 1240
 Db 366 SerIleAlaIleIleIleIleAsnSerAsnThrThrAspGlnGluAspAspTyArgPile 385
 QY 1241 GCCCTCAATGGGCTGTGCCAAGCCCTGACCCCTGTCCGCTCAATCCACCTGCTGCTC 1300
 Db 386 AlaLeuMetArgLeuSerLysProLeuThrLeuSerAlaHisIleHisProAlaCysLeu 405
 QY 1301 CCCATGATGAGACAGACCTTACCTTCATAGACACCTGCTGATCAAGAGCTTGGCAG 1360
 Db 406 ProMetHisGlyGlnThrPheSerLeuAsnGlnThrCysTrpIleThrGlyPheGlyLys 425
 QY 1361 ACCAGGAGACAGATGACAGACATCCCTCTCCGAGGAGTGCAGTCAATCTCATC 1420
 Db 426 ThrArgGlnThrAspAspLysThrSerProPheLeuArgGluValGlnValAsnLeuIle 445
 QY 1421 GACTTCAAGAAATGCAATGACTCTGTGTATGACAGTTACCTTACCCCAAGGATATG 1480
 Db 446 AspPheLysLysCysAsnAspTyThrLeuValThrAspSerTyThrLeuThrProArgMetMet 465
 QY 1481 TGTGCTGGGACCTTGTGGGGGCAAGACTCTCTGCCAGGAGACACAGGGGGGCTCTT 1540
 Db 466 CysAlaGlyAspLeuArgGlyGlyArgAspSerCysGlnGlyAspSerGlyGlyProLeu 485
 QY 1541 GTCTGTGAGCAGAAACAACCGCTGTACTGAGAGTGTCAACAGCTGGGGCAGAGCTGT 1600
 Db 486 ValCysGlnGlnAsnAsnArgTrpTyThrLeuAlaGlyAlaThrSerTrpGlyThrGlyCys 505
 QY 1601 GGCCAGAGAAACAACCTGTGTGTATACACCAAGTGAACAGAGTCTTCCCTGATTTAC 1660
 Db 506 GlyGlnArgAsnLysProGlyValTyThrThrValThrGluValLeuProTrpIleTy 525
 QY 1661 AGCAAGTGAAGAGCGAGGTGCGATTCAAGAAATCC 1696
 Db 526 SerLysMetGlnSerGlnValArgPheArgLysSer 537

Search completed: September 16, 2006, 02:41:55
 Job time : 327.5 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: September 16, 2006, 02:42:09 i Search time 10 Seconds
(without alignments)
3704.748 Million cell updates/sec

Title: US-10-806-370-11
Perfect score: 3333
Sequence: 1 cccagcagaccatgcagagagg9.....ggctcgtcgtcagcagaa 1748

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 256596 segs, 70647373 residues
Total number of hits satisfying chosen parameters: 513192

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10806370/runat.15092006.105651.10812/app_query.fasta.1
-DB=Published Applications AA New -OPWT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOEXT=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abss02p
-USER=US10806370 @CGN 1.1 68 @runat.15092006.105651.10812 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -IONGLOG -DEV TIMEOUT=120
-MAEN TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:
1: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US10_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US11_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	729.5	21.9	453	6	US-10-196-749-64 Sequence 64, Appli
2	726	21.8	454	7	US-11-400-825-2 Sequence 2, Appli
3	694	20.8	432	6	US-10-196-749-330 Sequence 330, Appli
4	694	20.8	432	7	US-11-101-316-112 Sequence 112, App
5	694	20.8	432	7	US-11-376-673-112 Sequence 112, App
6	608.5	18.3	416	7	US-11-208-257-8 Sequence 8, Appli
7	543.5	16.3	305	7	US-11-293-697-3002 Sequence 3002, Ap
8	532.5	16.0	853	6	US-10-530-187-242 Sequence 242, App
9	532.5	16.0	855	7	US-11-254-185-2 Sequence 2, Appli

10	532.5	16.0	855	7	US-11-253-869-2	Sequence 2, Appli
11	532.5	16.0	855	7	US-11-303-608-1	Sequence 1, Appli
12	525.5	15.8	255	7	US-11-254-185-6	Sequence 6, Appli
13	525.5	15.8	255	7	US-11-254-185-38	Sequence 38, Appli
14	525.5	15.8	255	7	US-11-253-869-6	Sequence 6, Appli
15	525.5	15.8	255	7	US-11-253-869-38	Sequence 38, Appli
16	523	15.7	531	6	US-10-522-668-2	Sequence 2, Appli
17	522	15.7	812	7	US-11-318-939-7	Sequence 7, Appli
18	520.5	15.6	423	6	US-10-196-749-320	Sequence 320, App
19	520.5	15.6	423	7	US-11-101-316-106	Sequence 106, App
20	520.5	15.6	423	7	US-11-376-673-106	Sequence 106, App
21	515.5	15.5	311	7	US-11-327-690-41	Sequence 41, Appli
22	510.5	15.3	343	7	US-11-359-554-3	Sequence 3, Appli
23	506	15.2	338	7	US-11-318-939-10	Sequence 10, Appli
24	505	15.2	276	7	US-11-327-690-35	Sequence 35, Appli
25	502.5	15.1	812	7	US-11-318-939-12	Sequence 12, Appli
26	498.5	15.0	339	7	US-11-404-745-1	Sequence 1, Appli
27	493.5	14.8	790	7	US-11-350-703-1	Sequence 1, Appli
28	493.5	14.8	791	7	US-11-318-939-6	Sequence 6, Appli
29	493	14.8	333	7	US-11-318-939-8	Sequence 8, Appli
30	492.5	14.8	810	7	US-11-431-663-2	Sequence 2, Appli
31	492.5	14.8	810	7	US-11-431-663-12	Sequence 12, Appli
32	486.5	14.6	790	7	US-11-318-939-13	Sequence 13, Appli
33	485.5	14.6	311	7	US-11-359-858-2	Sequence 2, Appli
34	480.5	14.4	262	7	US-11-359-554-2	Sequence 2, Appli
35	479	14.4	245	7	US-11-254-185-36	Sequence 36, Appli
36	479	14.4	245	7	US-11-253-869-36	Sequence 36, Appli
37	479	14.4	275	6	US-10-530-798-24	Sequence 24, Appli
38	478	14.3	275	6	US-10-530-798-25	Sequence 25, Appli
39	476.5	14.3	272	7	US-11-327-690-37	Sequence 37, Appli
40	475.5	14.3	272	7	US-11-359-554-6	Sequence 6, Appli
41	474.5	14.2	655	7	US-11-242-617-1	Sequence 1, Appli
42	473	14.2	275	6	US-10-530-798-26	Sequence 26, Appli
43	472.5	14.1	810	7	US-11-318-939-11	Sequence 11, Appli
44	470.5	14.1	272	7	US-11-327-690-40	Sequence 40, Appli
45	465	14.0	275	6	US-10-530-798-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-10-196-749-64
Sequence 64, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Guney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT FILING DATE: 2002-07-16
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: US/10/196, 749
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24


```

Db      55 LeuGlyIlelleAlaLeuIleAlaLeuAlaIleGlyLeuGlyIleHisPheAspCys 74
QY      560 TGGCAGGGCCACACAGGATCAGGTACAGAGAGAGAGAGAGAGCTGCCACGACGCT 619
Db      75 -----SerGlySerTyraArgCysArgSerSerPheLysCysIleGluLeuIle 90
QY      620 GTTCGCTGTGACGGGGGTGTGTGACCTGACAGCTGACAGAGTGAAGCTGGGCTGTGAG 679
Db      91 ThrArgCysAspGlyValSerAspCysGlyValAspGlyValArgCysValArg 110
QY      680 TTTCAGTGGAGCAAGCTCTGTTAAATCTACTGCTGGCTCTCCATCAGTGGCTTCCC 739
Db      111 ValGlyGlyIleAlaValIleGluInValPheThrAlaIleAsp-----TrpLysThr 128
QY      740 ATCTGTACGACAGCACTGGATGACTCTCTACTCAGAGAAAGACCTGCCAGCTGGGCTTC 799
Db      129 MetCysSerAspAspTrpLysGlyHisTyraAlaAsnValAlaCysAlaGluLeuGlyPhe 148
QY      800 GAGAGTCTCACCGGACACCGGAGGTTGCC-----CACAGG 835
Db      149 ProSerTyraValSerSerAspAsnLeuArgValSerSerLeuGlyGluPheArgGlu 168
QY      836 GATTTTGGCCAACAGCTTTCATCTTGGATACAC-----TCCACCATCCAGAAAGC 889
Db      169 GluPheValSerIleAspHisLeuLeuProAspAspLysValThrAlaLeuHisIleSer 188
QY      890 CTCACAC---AGGCTGAATGCCCTTCCACGGGATATCTCCCTCCAGTGTCCACTGC 946
Db      189 ValTyraValArgGlyGlyCysAlaSerGlyHisValValThrLeuGluCysThrAlaCys 208
QY      947 GGA---CTGAGGGGCATACCGGGCGGATGCTGGAGGGGGGCGCTGGCATAGCAAG 1003
Db      209 GlyHisArgArgGlyTyraSerSerArgIleValGlyIleAsnMetSerLeuLeuSerGlu 228
QY      1004 TGGCCTTGGCAAGTGAAGTGTGACCTTGGGACACACCATCTGTGTGAGGACAGCTCAT 1063
Db      229 TrpProTrpGlnAlaSerIleuGluInPheGluGlyTyraHisLeuGlyGlySerValIle 248
QY      1064 GAGGCCAGTGGGTGTGACTGCTGCCGCGCACTGCTTCTTCTGTAACCCGGAGAAAGTCTG 1123
Db      249 ThrProLeuTrpIleIleThrAlaAlaHisCysValTyraAspLeu-----TyraLeuPro 266
QY      1124 GAGGGCTGGAAGGTGTAGCGGGGACACAGACAGCACTGCCAGCTGGTGAAGGAGCTCC 1183
Db      267 LysSerThrThrIleGluInValGlyLeuValSerLeuLeuAspAsnProAlaProSerHis 286
QY      1184 ATTGCCAG---ATCATCATCATCAGCAATTAACCCGATGAGAGAGCACTATGATCATC 1240
Db      287 LeuValGlyIleValIleValIleSerSerTyraTyraProLysArgLeuGlyAsnAspIle 306
QY      1241 GCCCTCATGCGGCTGTGCAAGCCCTTACCTGCTCCCTCATCATCAGCTGCTGCTC 1300
Db      307 AlaLeuMetLysLeuAlaGlyProLeuThrPheAsnGluMetIleGluInProValCysLeu 326
QY      1301 CCCATGATGAGCAGACACTTAAGCTCATAGACAGCTGATGATCAGGCTTGGGAG 1360
Db      327 ProAsnSerGluAsnPheProAspGlyLysValCysTrpTrpSerGlyTrpGlyAla 346
QY      1361 ACCGAGGAGACAGATGACAGACATCCCTCTCCGAGAGTGCAGTCAATCATC 1420
Db      347 ThrGluAspGlyGlyAsp---AlaSerProValLeuAsnHisAlaIleAlaProLeuIle 365
QY      1421 GACTTCAGAAATGCAATGACTACTGTGTATGACAGTTACTTACCCCAAGATGATG 1480
Db      366 SerAsnLysIleCysAsnHisArgAspValTyraIleGlyIleIleSerProSerMetLeu 385
QY      1481 TGTCTGGGAGCACTTCTGGGGGAGAGACTCTGCCAGGAGAGACAGCGGGGCTCTT 1540
Db      386 CysAlaIleTyraLeuThrIleGlyIleValAspSerCysGluGlyAspSerGlyIleProLeu 405
QY      1541 GTCTGTGAGCAGAAACACCGCTGTACTCTGGCAGGTGTACACAGCTGGGAGCAGCTGT 1600
Db      406 ValCysGluGluArgArgLeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys 425

```

```

QY      1601 GGCCAGAAACAAACCTGTGTGTACACCAAGTGCAGAAAGTTCCTCGATTTAC 1660
Db      426 AlaGluValAsnLysProGlyValTyraArgValThrSerPheLeuAspTrpIleHis 445
QY      1661 AGCAAGATGAGAGCGAGGTGCGA 1684
Db      446 GluGluMetGluArgAspLeuLys 453

RESULT 3
US-10-196-749-330
; Sequence 330, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P9430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 330
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-330

Alignment Scores:
Pred. No.: 4,57e-34 Length: 432
Score: 694.00 Matches: 151
Percent Similarity: 51.7% Conservative: 73
Best Local Similarity: 34.9% Mismatches: 161
Query Match: 20.8% Indels: 48
DB: 6 Gaps: 11

US-10-806-370-11 (1-1748) x US-10-196-749-330 (1-432)
QY      455 CTGCGCAAGTTCACCTGGCGGAGGCGCAGAGAGCTACCGTATCGGAGTGTGCTC 514
Db      24 ILeProMetGluThrPheArg-----LysValGlyIleProIleIleAlaLeuLeu 41
QY      515 CTCCTCATTCCTCGTGTGTGCTTCATCATCTC-----TTCAG 556

```

```

Db      42 SerLeuAlaSerIleIleValValIleuIleValIleuAspIleValIleuVal 61
QY      557 TTCTGGAGAGGCGCCACAGGATGATGATACAGAGAGGAGGAGACTGTCACCAAC 616
Db      62 PheLeuCySgIyGln-----ProLeuHis 69
QY      617 GCTGTT-----CGCTGTACGGGTGTGTGACTGTCAAGCTGAAGAGTGACGAG 664
Db      70 PheIleProArgIySgIleuCyAspIyGluLeuAspIySgProLeuGluAspIy 89
QY      665 CTGGGCTGC-----GTGAGTTTGACTGGGAC 691
Db      90 GluHisCySValIySerPheProGluGlyProAlaValAlaValArgLeuSerIyAsp 109
QY      692 AAGCTCTGCTTAATCTACTGTGGGTCTCCCATGAGGCTTCCCATCTGAGAGC 751
Db      110 ArgSerThrIleuGlnValIleuAspSerAlaThrIleuIleuTrpPheSerAlaCySPhAsp 129
QY      752 AACTGATGATCTCTACTCTGAGAGAGACCTGCAGAGCTGGGTTCGAGAGTGTAC 811
Db      130 AsnThrThrIleuAlaIleuAlaGluThrAlaCySArgIleuMetGlyTrsSer----- 146
QY      812 CGGACAAACCGAGTTGCCACAGGATTTGCCAACAGCTTCTCAATCTTGAGATACAC 871
Db      147 ArgAlaValGluIleGlyProAspGlnAspLeuAspValValGluIleThrIleuAsnSer 166
QY      872 TCCACCATTCAGAGAAAGCTCCACAGGCTGTAAGTCCCTCCAGCGGTATATCTCCCTC 931
Db      167 GluHisIleuArgIleuValIleuAsnSerSerGlyProCySLeuSerGlySerLeuValSerIleu 186
QY      932 CAGTGTCCCACTGCTGAGTGTGAGGCGCATGACCGGCGGATCTGTGGAGAGGCGCTGACC 991
Db      187 HisCySLeuAlaCySgIyIySerIleuIleuThrProArgValValGlyIyGluAla 206
QY      992 TCGATATGCAAGTGGCTTGGCAAGTGAAGTGTGCACTTGGCAGCACCACATCTGTGGA 1051
Db      207 SerValAspSerTrpProIleuGlnValSerIleuIleuTrpAspIySgIleuValCySgIy 226
QY      1052 GGCAGCGCTCATGTAGCGCCAGTGGGTGTCTACGTGCGGCGCATCTCTTCTGTGACCGCG 1111
Db      227 GlySerIleuAspProIleuTrpValIleuThrAlaIleuHisCySPh-----ArgIyS 244
QY      1112 GAGAGAGTCTCGAGAGGCTGTGAGATGTACCGGCGGACCAAGCACTGACCAAGTGTGCT 1171
Db      245 HisThrAspValPheAsnTrpIyValArgAlaGlySerAspIySgIySerPhePro 264
QY      1172 GAGCAGCTTCATTTGCCAGATCATCTC-----AACAGCAATTACCGCATGAG 1222
Db      265 SerLeuAla--ValAlaIySgIleIleIleIleIleIleIleIleuAsnPrometIy----- 280
QY      1223 GAGAGCACTATGATGACGCGCTCATGCGGCTGTCCAGCGCCGATCGCTGCGCTCAC 1282
Db      281 ProIyAspAsnAspIleAlaIleuMetIyLeuGlnPheProLeuThrPheSerGlyThr 300
QY      1283 ATCCACCTGCTGTCTCTCCCATCATGACAGACCTTTAGCTTCAATGAGACCTGTGCG 1342
Db      301 ValArgProIleuCySLeuProPhePheAspGluIleuThrProAlaThrProLeuTrp 320
QY      1343 ATCAGAGCTTTGGCAAGACGAGGAGACAGATGACAAGACATCCCTTCTCGCGGAG 1402
Db      321 IleIleGlyTrpGlyPheThrIySgIleuAsnGlyIyIySgMetSerAspIleLeuIleuGln 340
QY      1403 GTGAGGTCATCTCATGATCTTCAAGAAATGACATGACTACTGTGCTATGACAGTTAC 1462
Db      341 AlaSerValGlnValIleAspSerThrArgCySAsnAlaAspAspAlaIyTrpGlnGly 360
QY      1463 CTTACCCCAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1522
Db      361 ValThrGluIySgMetMetCySAlaGlyIleProGluGlyIyValAspThrCySgIleu 380
QY      1523 GACAGCGGGGGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1582

```

```

Db      381 AspSerGlyIyProLeuMet---TyGlnSerAspGlnTrpHisValValIleVal 399
QY      1583 AGCTGGGGGACAGCTGTGTGCGAGAGAAACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1642
Db      400 SerTrpIyTrpIyGlyCySgIyGlyProSerThrProIyValIyTrpIySgValSerAla 419
QY      1643 GTTCTTCCTCGATTTTACAGCAAGATGAGAGCGAGGTG 1681
Db      420 TyIleuAsnTrpIleuTrpIleuValIleuValIleuValIleuValIleuValIleu 432

RESULT 4
US-11-101-316-112
: Sequence 112, Application US/11101316
: Publication No. US20060099657A1
: GENERAL INFORMATION:
: APPLICANT: Goddard, Audrey
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Guiney, Austin L.
: TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
: TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
: FILE REFERENCE: P323OR1C17C1
: CURRENT APPLICATION NUMBER: US/11/101,316
: PRIOR FILING DATE: 2005-04-06
: PRIOR APPLICATION NUMBER: 10/063526
: PRIOR FILING DATE: 2002-05-03
: PRIOR APPLICATION NUMBER: 10/006867
: PRIOR FILING DATE: 2001-12-06
: PRIOR APPLICATION NUMBER: PCT/US00/23328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: 09/380137
: PRIOR FILING DATE: 1999-08-25
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: NUMBER OF SEQ ID NOS: 170
: SEQ ID NO 112
: LENGTH: 432
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-101-316-112

Alignment Scores:
Pred. No.: 4,57e-34 Length: 432
Score: 694.00 Matches: 151
Percent Similarity: 51.7% Conservative: 73
Best Local Similarity: 34.9% Mismatches: 161
Query Match: 20.8% Indels: 48
DB: 7 Gaps: 11

US-10-806-370-11 (1-1748) x US-11-101-316-112 (1-432)
QY      455 CTGCGCAAGTTTACCTTGGCGGAGGCGCCAGAGCACTGATCGGAGTGTGCTCTC 514
Db      24 IleProMetGlnThrPheArg-----IySgValGlyIleProIleIleIleAlaLeuLeu 41
QY      515 CTCTCATTTGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
Db      42 SerLeuAlaSerIleIleIleValValIleuIleValIleuAspIyTrpIyTrp 61
QY      557 TTCTGGAGAGGCGCCACAGGATGATGATACAGAGAGGAGAGAGCTGTCCCAACGAC 616
Db      62 PheLeuCySgIyGln-----ProLeuHis 69
QY      617 GCTGTT-----CGCTGTACGGGTGTGTGACTGTCAAGCTGAAGAGTGACGAG 664
Db      70 PheIleProArgIySgIleuCyAspGlyGluLeuAspCySProLeuGluIyGluAspGlu 89
QY      665 CTGGGCTGC-----GTGAGTTTGACTGGGAC 691
Db      90 GluHisCySValIySerPheProGluGlyProAlaValAlaValArgLeuSerIyAsp 109

```

```

QY 692 AGCTCTGCTTAATACTACTGAGTCCCTCCATCAGTGGCTTCCCATCTGTAGACAGC 751
Db 110 ArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAsp 129
QY 752 AACTGGAATGACTCTCTACAGAAAGAGCTGCAGCAGCTGGAGTTTCGAGAGTGTCTAC 811
Db 130 AsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMetGlyTrpSer----- 146
QY 812 CGGACAAACGAGGTTGGCCACAGGATTTTGGCAAGCTTTCATCTTGAGATACAC 871
Db 147 ArgAlaValGlnIleGlyProAspGlnAspLeuAspValValGlnIleThrGlnAsnSer 166
QY 872 TCACACATCCAGAAAGCTCCACAGCTGTGAATGCCCTTCCAGCGGTATATCTCCTC 931
Db 167 GlnGlnLeuArgMetArgAsnSerSerglyProCysLeuSerglySerLeuValSerLeu 186
QY 932 CAGTGTCCCATGCGGACTGAGGAGCCATGACCGGCGATGCTGGAGGAGGCGCTGGCC 991
Db 187 HisCysLeuAlaCysGlySerglySerLeuLeuThrProArgValValGlyGlnGlnAla 206
QY 992 TCGGATAGCAAGTGGCTTGGCAAGTGAAGTCTGCATCTGGCAACACCATCTGTGA 1051
Db 207 SerValAspSerTrpProTrpGlnValSerIleGlnTrpAspGlySglnIleValCysGly 226
QY 1052 GGCACGCTATTGACCGCCAGTGGTGTCTCATCTGCCGCCCATCTCTTCTTGACCCGG 1111
Db 227 GlySerIleLeuAspProHisStrpValLeuThrAlaIleHisCysPhe-----ArgLys 244
QY 1112 GAGAAAGTCTTGAGGAGTGAAGGTATAGCGGGGACAGCAACCTGCACAGATTGCTCT 1171
Db 245 HisThrAspValAlaPheAsnTrpLysValArgAlaGlySerAspLysGlySerPhePro 264
QY 1172 GAGGACACCTTCATCTGGCCAGATCATCATC-----AACAGCAATTACACCGATGAG 1222
Db 265 SerLeuAla-----ValAlaLysIleIleIleIleGlnPheAsnProMetLys----- 280
QY 1223 GAGGACGACTATGACATCGCCCTATATCGGCTGTCCAGCCCTGACCCCTGCTCCGCTAC 1282
Db 281 ProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerglyThr 300
QY 1283 ATCCACCTGCTGCTCCCTCCATGACAGACAGACCTTTAGCCCTCATGAGACCTGCTGG 1342
Db 301 ValArgProLecCysLeuProPhePheAspGlnIleLeuThrProAlaThrProLeuTrp 320
QY 1343 ATCAAGGCTTTTGGCAAGACAGGAGACAGATGACAGACATCCCTTCTCCGCGAG 1402
Db 321 IleIleGlyTrpGlyPheThrGlnLysGlnAsnGlyLysLysMetSerAspIleLeuLeuGln 340
QY 1403 GTGCAGGTCATCTCATGACCTTCAAGAAATGCAATGACTATCTTGTGTATGACAGTTAC 1462
Db 341 AlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGlnGlyLys 360
QY 1463 CTTAACCCCAAGAGATGATGTGTGCTGGGAGCCTTGCTGGGAGGAGAGACTCTGCAGGGA 1522
Db 361 ValThrGlnLysMetCysAlaGlyIleLeuProGlnIleGlyValAspThrCysGlnGly 380
QY 1523 GACAGCGGGGCGCTCTGTCTGTGACAGACAGAAACCGCTGTACTCTGGACGTGTACCC 1582
Db 381 AspSerglyLysProLeuMet-----TyrGlnSerAspGlnTrpHisValGlyIleVal 399
QY 1583 AGCTGGGGCAAGGCTGTGGCCAGAGAAACAACTGGTGTGTACACCAAGTACAGAA 1642
Db 400 SerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 419
QY 1643 GTTCTTCCCTGATTTACAGCAAGATGAGAGGAGGAGG 1681
Db 420 TyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 432

```

RESULT 5
 ; Sequence 112, Application US/11376673
 ; Publication No. US20060160186A1

```

; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gueney, Auecin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
; FILE REFERENCE: P3230R1C165C
; CURRENT APPLICATION NUMBER: US/11/376, 673
; CURRENT FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: 10/063742
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 112
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-11-376-673-112

Alignment Scores:
Pred. No.: 4,57e-34 Length: 432
Score: 694.00 Matches: 151
Percent Similarity: 51.7% Conservative: 73
Best Local Similarity: 34.9% Mismatches: 161
Query Match: 20.8% Indels: 48
DB: Gaps: 11

US-10-806-370-11 (1-1748) x US-11-376-673-112 (1-432)
QY 455 CTGCCCAAGTTCACTGCGGCGGAGGCGCAGAACAGCTACCGCATCGGAGTGGCTC 514
Db 24 IleProMetGlnThrPheArg-----LysValGlyIleProIleIleIleAlaLeu 41
QY 515 CTCTCATTCGCTGGTGTGTTGCTCATCATCTC-----TTCCAG 556
Db 42 SerLeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTrpTyr 61
QY 557 TTCTGGCAGGCGCACACAGGAGTACAGTACAGAGCAGAGAGCTGTCCCAACAC 616
Db 62 PheLeuCysGlyGln-----CGCTGTACCGGGGTGTGTGACAGCTGAGAGTGAAGAG 664
QY 617 GCTGTT-----CGCTGTACCGGGGTGTGTGACAGCTGAGAGTGAAGAGTGAAGAG 664
Db 70 PheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCysAspProLeuGlyGlnAspGln 89
QY 665 CTGGGCTGC-----GTGAGGTTTGACTGGGAC 691
Db 90 GlnHisCysValLysSerPheProGlnGlyProAlaValAlaValArgLeuSerLysAsp 109
QY 692 AAGTCTCTGTTAAATCTACTGTGGCTCTCCCATGAGTGGCTTCCCATCTGTAGCAGC 751
Db 110 ArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAsp 129
QY 752 AACTGGAATGACTCTCTACAGAAAGAGCTGCAGCAGCTGGAGTTTCGAGAGTGTCTAC 811
Db 130 AsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMetGlyTrpSer----- 146
QY 812 CGGACAAACGAGGTTGGCCACAGGATTTTGGCAAGCTTTCATCTTGAGATACAC 871
Db 147 ArgAlaValGlnIleGlyProAspGlnAspLeuAspValValGlnIleThrGlnAsnSer 166
QY 872 TCACACATCCAGAAAGCTCCACAGCTGTGAATGCCCTTCCAGCGGTATATCTCCTC 931
Db 167 GlnGlnLeuArgMetArgAsnSerSerglyProCysLeuSerglySerLeuValSerLeu 186
QY 932 CAGTGTCCCATGCGGACTGAGGAGCCATGACCGGCGATGCTGGAGGAGGCGCTGGCC 991

```

```

Db      187 HisCysLeuAlaCysGlyLysSerLeuYerThrProArgValAlaGlyGlyGluAla 206
Qy      992 TCGGATAGCAAGTGGCCCTGGCAAGTGTGACCTGGCACCACCCATCTGTGGA 1051
Db      207 SerValAspSerTrpProTyrGlnValSerIleGlnTyrAspLysGlnIleValCysGly 226
Qy      1052 GGCAGCCTCATTTAGCCGCAAGTGGGTCTCACTGCGCCGCACTGCTTCTTGTAACCGG 1111
Db      227 GlySerIleLeuAspProHleTrpValIleThrAlaAlaHisCysPhe-----ArgLys 244
Qy      1112 GAGAAAGTCTCGAAGGCTGGAAGGTGTACCGGGGACACCAACCTGGACCAAGTTGCT 1171
Db      245 HisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhePro 264
Qy      1172 GAGGACGCTTCATGCGGAGATCATCATC-----AACAGCAATTACACGAGTAG 1222
Db      265 SerLeuAla--ValAlaLysValIleIleIleIleGluPheAsnProMetTyr----- 280
Qy      1223 GAGGACGACTATGACATCGCCCTCATGCGGCTGTCAAGCCCTGACCCCTGCTGCAC 1282
Db      281 ProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThr 300
Qy      1283 ATCCACCTGCTGCTTCCCTCCCATGACATGACACACTTACCTTCAATGAGACCTGCTGG 1342
Db      301 ValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThrProLeuTrp 320
Qy      1343 ATCAGACGCTTTGGCAAGACCGAGGAGACAGATGACAGACATCCCTTCCTCGCGAG 1402
Db      321 IleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuGln 340
Qy      1403 GTGAGGCTCATCTCATGACCTTCAAGAAATGACATGACTCTTGCTATGACAGTTAC 1462
Db      341 AlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyLys 360
Qy      1463 CTTATCCCAAGATGATGTGTGTGGGACCTTGTGGGGGACAGACTCTGCCAGGGA 1522
Db      361 ValThrGluLysMetCysAlaGlyIleProGluGlyValAspThrCysGlnGly 380
Qy      1523 GACGCGGGGGGCTCTTGTGTGTGAGAGAAACCGCTGTACCTGGGACGGTGTACC 1582
Db      381 AspSerGlyGlyProLeuMet--TyrGlnSerAspGlnTrpHisValAlaGlyIleVal 399
Qy      1583 AGCTGGGACAGGCTGTGGCCAGAGAAACAACTGGTGTGTACACCAAGTACAGAA 1642
Db      400 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 419
Qy      1643 GTTCTTCCCTGGATTTTACAGCAATGAGACAGCAGGTG 1681
Db      420 TyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 432

```

RESULT 6

```

US-11-208-257--8
/ Sequence 8, Application US/11208257
/ Publication No. US20060101531A1
/ GENERAL INFORMATION:
/ APPLICANT: Vaseiouhin et al., Valeri
/ TITLE OF INVENTION: ANIMAL MODELS OF CANCER DEVELOPMENT AND METASTASIS
/ FILE REFERENCE: 14539A-008510US
/ CURRENT APPLICATION NUMBER: US/11/208,257
/ PRIOR FILING DATE: 2005-08-19
/ PRIOR APPLICATION NUMBER: US 60/603,043
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 8
/ LENGTH: 416
/ TYPE: PRT
/ ORGANISM: mouse
US-11-208-257-8

```

Alignment Scores: 5.75e-29 Length: 416

```

Score: 608.50      Matches: 140
Percent Similarity: 52.6%      Conservative: 60
Best Local Similarity: 36.8%      Mismatches: 135
Query Match: 18.3%      Indels: 45
DB: 7      Gaps: 12

```

US-10-806-370-11 (1-1748) x US-11-208-257-8 (1-416)

```

Qy      635 GTGGTGACGTCGAACGAGAGAGAGAGAG-----CTGGGCTGTGGAGTTTGACTGG 688
Db      38 IleValThrIleLeuLeuGlnSerAspGlnGluProLeuTyrGlnValGlnLeuSerPro 57
Qy      689 GACAAGTCTGTCTTAAATCTACTGTGGTCTCCCATGAGTGGCTTCCCATGTAGC 748
Db      58 GlyAspSerThrArgLeuAlaValLeuAspLysThrGlnGlyThrTrpArgLeuLeuCysSer 77
Qy      749 AGCACTGGAATGACTCTTACTCAGAGAAGACCTGCCAGACGCTGGTTTCGAGAGTCT 808
Db      78 SerArgSerAsnAlaArgValAlaGlyLeuGlyCysGlnGluMetGlyPheLeuArgAla 97
Qy      809 CACCGGACAAACGAGGTTGCCACAGGATTTT-----GCCACAGC 850
Db      98 -----IleAlaHisSerGluLeuAspValArgThrAlaGlyAlaAsnGly 112
Qy      851 TTCTCA-----ATCTTGAGATPACAATCCACCATCCAGAA 886
Db      113 ThrSerGlyPhePheCysValAspGlnGlyGlyLeuProLeuAlaGlnArgLeuLeuAsp 132
Qy      887 AGCTCCACAGGTGTGAATGCTCCCATCCAGGATATTCCTCCCATGTTCCTCACTGC 946
Db      133 ValIleSerValCysAspCysProArgGlyArgPheLeuThrAlaThrCysGlnAspCys 152
Qy      947 GACTGAGGCGCAATG---ACCGGCGGATGATGGAGAGGCGCTGCGCTGAGATAGCAAG 1003
Db      153 GlyArgArgLysLeuProValAspArgIleValGlyGlnIleAspSerLeuGlyArg 172
Qy      1004 TGGCTTGGCAAGTGAAGTGTGACCTTGGACCAACCCACATCTGTGGAGCAGCCTATT 1063
Db      173 TrpProTrpGlnValSerLeuArgTyrAspGlyThrHisLeuGlyGlySerLeuLeu 192
Qy      1064 GACGCCAGTGGGTGTCTCATCTGCCCGCCACTGCTTCTTGTGACCCGGGAGAAAGTCTCT 1123
Db      193 SerGlyAspTrpValLeuThrAlaAlaHisCys---PheProGluArgAsnArgValLeu 211
Qy      1124 GAGGCTGGAAGGTGTACGGGCGC-----ACCGCAACCTGCACAGTTCCT 1171
Db      212 SerArgTrpArgValPheAlaGlyAlaValAlaArgThrSer-----Pro 226
Qy      1172 GAGGCAGCC-----TTCATTGCCGAGATCATCAACAGCAATTAC----- 1213
Db      227 HisAlaValAlaGlnLeuGlyValAlaValIleTyrHisGlyGlyTyrLeuProPheArg 246
Qy      1214 -----ACCGATGAGAGAGACACTATGACATGCCCTCATGCGGCTGTCCAGCCCTG 1267
Db      247 AspProThrIleAspGlnAsnSerAsnAspIleAlaLeuValHisLeuSerSerLeu 266
Qy      1268 ACCCTGCGCCGTCACATCCACCTGCTTGTGCTCCCATGATGAGACAGACTTATAGCTC 1327
Db      267 ProLeuThrGlnTyrIleGlnProValCysLeuProAlaAlaGlyGlnAlaLeuValAsp 286
Qy      1328 AATGAGACTGCTGTGATCAGAGCTTTGGCAAGACCGAGGAGACAGATGACAGACATCC 1387
Db      287 GlyLysValCysThrValThrGlyTyrGlyAsnThr---GlnPheTyrGlyGlnAla 305
Qy      1388 CCTTCTCCGGGAGGTCAGGTCAATCTCATGACTTCAAGAAATGCAATGACTACTTG 1447
Db      306 MetValLeuGlnGlnAlaArgValProIleIleSerAsnGlnValCysAsnSerProAsp 325
Qy      1448 GTCTATGACAGATTAACCTTACCCCAAGATGATGTGTGGGACCTTCTGTTGGGAGAGA 1507
Db      326 PheTyrGlyAsnGlnIleLysProLysMetPheCysAlaGlyTyrProGluGlyGlyIle 345
Qy      1508 GACTCTGTCCAGGAGACAGCGGGGCGCTTGTGTGTGAGCAGAAC----- 1555

```



```

Db      346 AspAlaCysGlnGlyAspSerGlyGlyProPheValCysGlnAspSerIleSerGlyThr 365
QY      1556 AACCGCTGTAAGCTGGCAGGTGTACACAGCTGGCGCACAGCTGTGGCCAGAAACAA 1615
Db      366 SerArgTyrArgLeuGlyGlyIleValSerTyrGlyThrGlyCysAlaLeuValArgGly 385
QY      1616 CCGGAGTGTACACCAAGTACGAAAGTCTTCCCTGATTTACAGCAGATGAGAGC 1675
Db      386 ProGlyValTyrThrIleValIleAspPheArgGlnTyrIlePheValIleValIleVal 405

RESULT 7
US-11-293-697-3002
; Sequence 3002, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3002
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3002

Alignment Scores:
Pred. No.: 4,29e-25 Length: 305
Score: 543.50 Matches: 121
Percent Similarity: 54.4% Conservative: 39
Best Local Similarity: 41.2% Mismatches: 107
Query Match: 16.3% Indels: 28
DB: Gaps: 9

US-10-806-370-11 (1-1748) x US-11-293-697-3002 (1-305)
QY      873 CCACATTCACGAAAGCTCCACAGGTGTGATGCCCTCCAGCGGTATATCTCCCTCC 932
Db      8 ProProArgAlaGlyProSerLeuGlySerGlyAspLeu-----GlySerProLeu 25
QY      933 AGTGTCCCA-----TCCGAGCTAGAGGCCATG---ACCGG 967
Db      26 SerProProAlaAspProCysProThrAspCysGlyArgArgLeuProValAsp 45
QY      968 CGGATGCTGGAGGGCGCTGGCGCTCGGATACGAGTGGCTTGGCAAGTGAATCTGCAC 1027
Db      46 ArgIleValGlyGlyArgPheThrSerLeuGlyArgTyrProTyrGlnValSerLeuArg 65
QY      1028 TTGGGACACCAACATCTGTGAGGACGCTGATTTAGCCCAAGTGGGTGCTCATGCGC 1087
Db      66 TyrAspGlyAlaIleValLeuGlyGlySerLeuSerGlyAspTyrValLeuIleAla 85
QY      1088 GCCCACTGCTTCTTCTGTGACCGCGGAGAGGCTCTGAGGAGCTGAGAGTGTACGCGGCG 1147
Db      86 AlaIleCys---PheProGlnArgGlnValIleLeuSerArgTyrArgValPheAlaGly 104
QY      1148 ACC-----AGCAACTGACACCACTGCTGAGGACCTTCATGCGCAGATCATC 1198
Db      105 AlaValAlaGlnIleAspProIleGlyLeu-----GlnLeuGlyValGlnIleValIle 122
QY      1199 ATCAACAGCATTAAC-----ACCGATGAGGAGGAGCACTATGACATC 1240
Db      123 TyrHisGlyGlyTyrLeuProPheArgAspProAsnSerGlnGlnIleValIleValIle 142
QY      1241 GCGCTTACGCGGCTGTCGAAGCCCTGACCTGTGCTGCTCATCATCACCCTGCTGCTTC 1300
Db      143 AlaLeuValHisIleuSerSerProLeuProLeuThrGlnTyrIleGlnProValCysLeu 162
```

```

QY      1301 CCCATCATGACAGACACTTTAGCCTCAATGAGACCTGTGATCACAGGCTTTGGCAAG 1360
Db      163 ProAlaIleGlyGlnIleValLeuValAspGlyGlyIleCysThrValThrGlyTyrGlyAsn 182
QY      1361 ACCGAGAGACAGATGACAGACATCCCTTCTCCGAGAGGTGACGTAATTCATC 1420
Db      183 ThrGlnTyrTyrGlyGlnIleValIleGly---ValLeuGlnGlnIleValArgValProIle 201
QY      1421 GACTTCAAGAAATGCAATGACTTGTGCTATGACAGTACTTACCCCAAGAGTATG 1480
Db      202 SerAsnAspValCysAsnGlyAlaAspPheTyrGlyAsnGlnIleValProValPhe 221
QY      1481 TGTGCTGGAGACCTTGTGAGGAGCAGACACTCTGCGCAGAGACAGCGGAGGCTCTT 1540
Db      222 CysAlaGlyTyrTyrProGlnGlyIleAspAlaCysGlnIleAspSerGlyGlyProPhe 241
QY      1541 GTCGTGAG-----CAGAACACCGCTGTACTGCGACAGGTGTACCAAGCTGG 1588
Db      242 ValCysGlnAspSerIleSerArgTyrProArgTyrArgLeuGlyIleValSerTyr 261
QY      1589 GGCACAGCTGTGGCCAGAGAAACAAACCTGTGTGTATACCAAGAGACAGAGATTCTT 1648
Db      262 GlyThrGlyCysAlaLeuValGlnIleValProGlyValTyrThrIleValSerAspPheArg 281
QY      1649 CCCTGATTTACAGCAAGATGAGAGCGAGGTGCGATTTCAGA 1690
Db      282 GluTyrIlePheGlnIleValIleValValGlySerArg 295

RESULT 8
US-10-530-187-242
; Sequence 242, Application US/10530187
; Publication No. US20060183120A1
; GENERAL INFORMATION:
; APPLICANT: THE, Bin Tean
; APPLICANT: TAKAHASHI, Masayuki
; TITLE OF INVENTION: Molecular Subclassification of Kidney Tumors and the Discovery of
; FILE REFERENCE: 28927-0014
; CURRENT APPLICATION NUMBER: US/10/530,187
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US2003/031476
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: US 60/415,775
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 242
; LENGTH: 853
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-530-187-242

Alignment Scores:
Pred. No.: 2,03e-24 Length: 853
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
DB: Gaps: 25

US-10-806-370-11 (1-1748) x US-10-530-187-242 (1-853)
QY      282 TTTCAGAGTCTCATCCGCGAGTCAATCATCCGCCAGGTGAGCTGAGTCAACTGCC 341
Db      349 PheAsnSerProTyrTyrProGlyHisTyr-----ProPro 360
QY      342 CAACAGAGTGAAGCTTTAGAGCAACACACATGGGAGGCTGTACCATCATCATCTTC 401
Db      361 AsnIleAspCysThr-----TyrAsnIleGlnValIleProAsnAsn--- 373
QY      402 CTGCGAGGTCAAGAC-----CAGCAACAGAGGCCCA 431
Db      374 -----GlnHisValValArgPheValPhePheTyrLeuLeuGlnProGlyVal 390
```

```

QY 432 CCAGGAGAGCCGAGTACGAGCCTGCCAAGTTCA----- 467
    |||
    |||
    |||
Db 391 Pro-----AlaGlyThrCysProLysAspTyrValGluIleAsnGlyGluLys 406
QY 468 -----CCTGGCCGGAGAGGCCAGA----- 485
    |||
    |||
    |||
Db 407 TyrCysGluIleuArgSerGlnPheValAlaThrSerAsnSerAsnLysIleThrValArg 426
QY 486 -----ACGAGTACCGGCTCATCGGGTGGCTGCTCCCTCATTTGCCCTGGTGG 533
    |||
    |||
    |||
Db 427 PheHisSerAspGlnSerTyrThrAspThrGlyPheLeuAlaGluTyrLeu----- 443
QY 534 TTTCGCTCATCATCTCTCTTCCAGTTTCGAGGGCCACACAGGATGACGATACAGAGC 593
    |||
    |||
    |||
Db 444 -----SerTyrAspSerSerAspProCysProGlyGlnPheThrCysArg 458
QY 594 AGAGGAGAGAGTGTCCAGAGCCGTGTGCTGTGACGGGGGTGTGACTGCACAGCTGA 653
    |||
    |||
    |||
Db 459 ThrGlyArg--CysIleArgGluLeuArgCysAspGlyTyrAlaAspCysThrAspH 478
QY 654 AGAGTGAAGAGTGGCTGGCTGAGGTTTACTGGGACAAGCTCTGCTTAAATCTACT 713
    |||
    |||
    |||
Db 478 IsSerAspGluLeuAsnCys-----Serc 486
QY 714 CTGGGCTCTCCCATCAG-----TGCGCTTC 737
    |||
    |||
    |||
Db 486 LysAspAlaGluIleGlnPheThrCysLysAsnLysPheCysLysPheProLeuPheThr 504
QY 738 CCATCTGTAGAGC-----AACTGGAATGACTCTTACTACAGAAAGACCTGC----- 784
    |||
    |||
    |||
Db 505 --ValCysAspSerValAsnAspCysGluYAspAsnSerAspGluGlnGlyCysSerCysP 524
QY 785 -----CAGAGCTGG 794
    |||
    |||
    |||
Db 524 LysIleGlnIlePheArgCysSerAsnGlyLysCysLeuSerLysSerGlnGlnCysAsnG 544
QY 795 GTTTCGAGAGTGTCCAGCCGACAAACGAGTTGCCACAGGATTTTGCCACAGCTTCT 854
    |||
    |||
    |||
Db 544 LysLysAspAspCysGluYAspGlySerAspGluAlaSerCysProLysValAlaValT 564
QY 855 CAATCTTGAGTACATCCATCCATCCAGAAAGCTCCACAGGCT-----G 902
    |||
    |||
    |||
Db 564 hTrCysThrLysHisThrTyrArgCysLeuAsnGlyLeuCysLeuSerLysGlyAsnProG 584
QY 903 AATGCCCTTCCACGGGATATCTCCCTCCAGTGTCC----- 940
    |||
    |||
    |||
Db 584 LysCysAspGlyLysGlu-----AspCysSerAspGlySerAspGluLysAspC 600
QY 941 --CACTGCGAGTGAAGGCCATGAC-----GGGCGGATCGTGGAGGGCGCTGGCT 992
    |||
    |||
    |||
Db 600 LysAspCysGluLeuArgSerPheThrArgGlnAlaArgValAlaGlyLysThrAspAla 620
QY 993 CGGATGACAGTGGCTTTGCGAAGTGAAGTGTGC--TTGCGACACACCCACATCTGTG 1049
    |||
    |||
    |||
Db 620 spGluGlyGlnTyrProTyrGlnValSerLeuHisAlaLeuGlyGlnGlyHisIleCysG 640
QY 1050 GAGGACAGCTCATTAAGCGCCAGTGGGTGCTCAGTGGCCACCTGCTCTCGTGAACC 1109
    |||
    |||
    |||
Db 640 LysAlaSerLeuIleSerProAsnTyrLeuValSerAlaAlaHisCysTyrIleAspAsp 660
QY 1110 GCGAGAGAGTCTCGAGAGG-----TGGAAGGTGTACGCG 1145
    |||
    |||
    |||
Db 660 rg-----GlyPheArgTyrSerAspProThrGlnTyrPheAlaPheLeuG 675
QY 1146 GCACCAAGACCTGCACAGTTGCTTGAAGCAGCCTCC-----A 1184
    |||
    |||
    |||
Db 675 Lys-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnGlnLysArgL 692
QY 1185 TTGCGGAGATCATGATCAACAGCAATTACACCGATGAGAGAGAGATGATGACATGCGCC 1244
    |||
    |||
    |||
Db 692 euLysArgIleIleSerHisPheProPheAsnAspPheThrPheAspTyrAspIleAla 712

```

```

QY 1245 TCATGCGGCTGTCCAGACCCCTGACCTGTCCGCTCATATCCACCTTGTGCTCCCCA 1304
    |||
    |||
    |||
Db 712 euLeuGluLeuGluLysProAlaGlyTyrSerSerMetValArgProIleCysLeuProA 732
QY 1305 TGCAATGACAGACCTTTAGCTCATAGACCTGTGATCAACAGGCTTGGCAAGCA 1364
    |||
    |||
    |||
Db 732 spAlaSerHisValPheProAlaGlyLysAlaIleTyrValHisGlyTyrPoleIsthr 751
QY 1365 GGGAGACAGATGACACAGATCCCTCTCTCCGGAGGTGACAGTCAATCTATGACT 1424
    |||
    |||
    |||
Db 752 --GlnTyrGlyGlyThrGlyAlaLeuIleLeuGlnLysGlyLysValIleAsnG 771
QY 1425 TCAGAAATGCATGACTACTTGTGTATGACAGTTACCTTACCCCAAGATGATGTGTG 1484
    |||
    |||
    |||
Db 771 LysThrCysGluAsnLeu-----ProGlnGlnIleThrProArgMetMetCysAv 789
QY 1485 CTGGGAGCCTTCTGGGGGGGAGAGACTCTGCCAGGAGGAGGAGGGGGCTT---G 1541
    |||
    |||
    |||
Db 789 AlGlyPheLeuSerGlyValAlaAspSerCysGlnGlyAspSerGlyLysProLeuSer 809
QY 1542 TCTGTGACAGAACCAACCGCTGTGACTCGGACAGGTGTACACAGCTGGGACAGCTGTG 1601
    |||
    |||
    |||
Db 809 euValGluAlaAspGlyArgIlePheGlnAlaGlyValAlaSerTyrGlyAspGlyCysA 829
QY 1602 GCCAGAGAAACAACTGTGTGTATCACCAAGTACAGAAATTTCTTCCCTGATT 1657
    |||
    |||
    |||
Db 829 LysGlnArgAsnLysProGlyValIleTyrThrArgLeuProLeuPheArgAspTyrIle 847

```

RESULT 9

US-11-254-185-2
 ; Sequence 2, Application US/11254185
 ; Publication No. US20060099625A1

; GENERAL INFORMATION:

; APPLICANT: CRAIK, CHARLES S.
 ; APPLICANT: TAKEUCHI, TOSHIIKO
 ; APPLICANT: SCHUMAN, MARC
 ; TITLE OF INVENTION: MT-SPI SERINE PROTEASE
 ; FILE REFERENCE: 28644-701.302
 ; CURRENT APPLICATION NUMBER: US/11/254,185
 ; CURRENT FILING DATE: 2005-10-18
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 2
 ; LENGTH: 855
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-11-254-185-2
 Alignment Scores:
 Pred. No.: 2,03e-24
 Score: 532.50
 Percent Similarity: 40.0%

Best Local Similarity: 29.1%
 Query Match: 16.0%
 Matches: 163
 Conserved: 61
 Mismatches: 175
 Indels: 161
 Gaps: 25

US-10-806-370-11 (1-1748) x US-11-254-185-2 (1-855)

```

QY 282 TTTCAGAGTCTTATCCGCGAGTCAATCCGCGAGTCAAGCTCGGTGACAACTTCCC 341
    |||
    |||
    |||
Db 351 PheAsnSerProTyrTyrProGlyHisTyr-----ProPro 362
QY 342 CAACCAAGTGTACTTGTATGAGCAACACAGTGGGGGTGTACCCATCGATCATCTC 401
    |||
    |||
    |||
Db 363 AsnIleAspCysThr-----TrpAsnIleGluValProAsnAsn--- 375
QY 402 CTGCGAGTGAAGC-----CAGCAACCGAGGCCA 431
    |||
    |||
    |||
Db 376 -----GlnHisValLysValArgPheLysPhePheTyrLeuLeuGlnProGlyVal 392
QY 432 CCAGGAGAGCCAGGATGACAGCTGCGCCAAAGTTCA----- 467
    |||
    |||
    |||
Db 393 Pro-----AlaGlyThrCysProLysAspTyrValGluIleAsnGlyGluLys 408

```

```

QY 468 ---CTGGCGGAGGCGACA----- 485
D 409 TyrCysGlyGluArgSerGlnPheValThrSerAsnSerAsnLysIleThrValArg 428
QY 486 -----AGACGTACCGCTCATCGGGTCGCTCTCCATTCATTCCTGGTGG 533
D 429 PheHisSerAspGlnSerThrAspThrGlyPheLeuAlaGluIleValArg 445
QY 534 TTTGGCTCATCATCTCTTCACAGTTTCGGACGGGCCACACAGGATCAGGTACAGAGAGC 593
D 446 -----SerTyrAspSerSerAspProCysProGlyGlnPheThrCysArg 460
QY 594 AGAGAGAGAGCTGCCACAGACCGCTGCTGTCAGAGGGGTGGTGGTGGTGGTGGTGGTGG 653
D 461 ThrGlyArg-CysIleArgGlySerGlnLeuArgCysAspGlyTyrPalalAspCysThrAsp 480
QY 654 AGAGTGAAGAGCTGGCTGCTGAGGTTTGACTGGGACAACTCTCTGCTTAAATCTACT 713
D 480 IAspSerAspGlnLeuAsnCys-----SerC 488
QY 714 CTGGCTCTCCATCAG-----TGCTTC 737
D 488 ysaAspAlaGlyHisGlnPheThrCysLysAsnLysPheCysLysProLeuPheTrp---- 506
QY 738 CCATCTGTAGAGC-----AACTGGAATGACTCTCTACTCAGAGAGACCTGC----- 784
D 507 --ValCysAspSerValAsnAspCysGlyAspAsnSerAspGlnGlnIleCysSerCysP 526
QY 785 -----CAGCAGCTGG 794
D 526 rolaGlnThrPheArgCysSerAsnGlyLysCysLeuSerLysSerGlnCysAsnG 546
QY 795 GTTTCGAGAGTCTCACCGGACAAACCGAGTTCCTCCACAGGAGATTGGCCACAGCTTCT 854
D 546 LysAspAspCysGlyAspGlySerAspGlnAlaSerCysProLysValAsnValArg 566
QY 855 CAATCTTGAGTATACCTCCACCATCCAGAAACCTCCACAGGCT-----G 902
D 566 hrcysThrLysHisThrTyrArgCysLeuAsnGlyLeuCysLeuSerLysGlyAsnProG 586
QY 903 AATGCCCTCCACGCGATATCTCCCTCCAGTGTCC----- 940
D 586 lncysAspGlyLysGlu-----AspCysSerAspGlySerAspGlnLysAspC 602
QY 941 --CACTCGGAGCTGAGGCGCATGAC-----GGCGGAGCTGAGGAGGCGCTGGCT 992
D 602 ysaAspCysGlyLeuArgSerPheThrArgGlnAlaArgValAlaGlyLysThrAspAla 622
QY 993 CGGATACAGAGTGGCTTTGGAAAGTGAAGTGGAC---TTGGGACACACCCACATCTGTG 1049
D 622 spGlnGlyGlnTrpProThrGlnValSerLeuHisAlaLeuGlyGlnGlyHisIleCysG 642
QY 1050 GAGGACAGCTCATTCAGCGCCAGTGGGTGCTCATCGCGCCACCTGCTTCTCTGACCC 1109
D 642 lylalaserleuIleSerProAsnTrpLeuValSerAlaHisCysIleThrIleAspAla 662
QY 1110 GGGAGAAAGTCTGAGAGGC-----TGAAAGGTGTAACGGG 1145
D 662 rg-----GlyPheArgTyrSerAspProThrGlnTrpAlaPheLeuG 677
QY 1146 GCACACAGCACTGACACAGTTGCTGAGGAGCGCTCC-----A 1184
D 677 ly-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnGlnArgL 694
QY 1185 TTGGCGAGATCATCATCAACGAAATTACCCGATGAGAGAGAGAGACTATGACTGGCC 1244
D 694 eulysargIleIleIleSerHisProPhePheAsnAspPheThrPheAspIleAlaIleAla 714
QY 1245 TCATGCGGCTGTCAAGCCCTGACCCCTGACCTCGCTGCATCATCCCTGCTTGGCTCCCC 1304
D 714 euleuGlnLeuGlnLysProAlaGlnLysSerSerMetValArgProIleCysLeuProA 734

```

```

QY 1305 TGCAATGACAGACCTTTAGCCTCAATGACCTGCTGATCAGAGCTTGGCAAGACA 1364
D 734 spAlaserHisValPheProAlaGlyLysAlaIleTrpValIleThrGlyTyrGlyHisThr- 753
QY 1365 GGGAGACAGATGACAAACATCCCTCTCCGGGAGGTGACAGTCAATCTCATGACT 1424
D 754 --GlnTyrGlyGlyThrGlyAlaLeuIleLeuGlnLysGlyGlnLysArgValIleAsnG 773
QY 1425 TCAAGAAATGCAATGACTACTGCTGCTATGACAGTTACCTTACCCCAAGATGATGTGTG 1484
D 773 lnhThrCysGluAsnLeu-----ProGlnGlnIleThrProArgPheThrCysV 791
QY 1485 CTGGGACCTTCGTGGGGGAGAGACTCTGCCAGGAGACAGCGGGGCGCTCTT---G 1541
D 791 alGlyPheLeuSerGlyGlyValAspSerCysGlnGlyAspSerGlyGlyProLeuSer 811
QY 1542 TCTGTGACAGAACACCGCTGTAAGTCTGCGAGGTGTCAACGCTGGGGCACAGCTGTG 1601
D 811 erylGlnAlaAspGlyArgIlePheGlnAlaGlyValIleSerTrpGlyAspGlyCysA 831
QY 1602 GCCAGAGAAACAAACCTGTGTGTGATACCAAGAGACAGAAAGTCTTCCCTGAGATT 1657
D 831 laglnArgAsnLysProGlyValTyrThrArgLeuProLeuPheArgAspTrpIle 849

RESULT 10
US-11-253-869-2
: Sequence 2, Application US/11253869
: Publication No. US20060104979A1
: GENERAL INFORMATION:
: APPLICANT: CRAIK, CHARLES S.
: APPLICANT: TAKEUCHI, TOSHIIKO
: TITLE OF INVENTION: MT-SP1 POLYNUCLEOTIDES AND POLYPEPTIDES
: FILE REFERENCE: 28644-701.303
: CURRENT APPLICATION NUMBER: US/11/253, 869
: NUMBER OF SEQ ID NOS: 83
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 2
: LENGTH: 855
: TYPE: PRT
: ORGANISM: Homo sapiens
US-11-253-869-2

Alignment Scores:
Pred. No.: 2,03e-24 Length: 855
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
Gaps: 25

US-10-806-370-11 (1-1748) x US-11-253-869-2 (1-855)
QY 282 TTTCCAGAGCTCTACCGGAGGTCAATCCGCGAGCTGAGCTCGGTGACAACTGCC 341
D 351 PheAsnSerProTyrTyrProGlyHisTyr-----ProPro 362
QY 342 CAACAGAGTGTACTTTGTAAGCAACACAGTGGGGGCTGTAACCATCGATCATCTC 401
D 363 AsnIleAspCysThr-----TrpAsnIleGluValProAsnAsn--- 375
QY 402 CTGGCAGGTGAGCAC-----CAGCAACCGGGCCA 431
D 376 -----GlnHisValLysValArgPheLysPheThrLeuLeuGlnProGlyVal 392
QY 432 CCAAGGAGAGCCCGAGTACAGCTGCGCCAAATTCA----- 467
D 393 Pro-----AlaGlyThrCysProLysAspTyrValGluIleAsnGlyLys 408
QY 468 ---CTGGCGGAGGCGGACA----- 485
D 409 TyrCysGlyGlnArgSerGlnPheValThrSerAsnSerAsnLysIleThrValArg 428

```

```

OY 486 -----ACAGTACCGCTCATCGGGTGGTGGTCTCTCTCATTTGCGCTGGTG 533
Db 429 PheHisSerArgInserTyrThrAspThrGlyPheLeuAlaGlyTyrLeu----- 445
OY 534 TTTCGCTCATCATCTCTTCCAGATTCTGGACGGGSCACACAGGATGACGATACAGAGC 593
Db 446 -----SerTyrAspSerSerAspProCysProGlyGlnPheThrCysArg 460
OY 594 AGAGGAGAGCTGTCCACAGACGCTGTGCTGGACGGGTGTGTGATGTGACAGCTGA 653
Db 461 ThrGlyArg-CysIleArgGlyLeuArgCysAspGlyTyrAlaAspCysThrAspH 480
OY 654 AGAGTGAAGAGCTGGCTGCTGAGAGGTTTGAAGTGAAGTCTGTGTTAAATCTACT 713
Db 480 IAspSerArgGlyLeuAsnCys-----SerC 488
OY 714 CTGGGCTCTCCATGAG-----TGCGCTTC 737
Db 488 ysaAspAlaGlyNIleGlnPheThrCysLeuAsnLysPheCysLeuProLeuPheTyr 506
OY 738 CCATCTGTAGAGC-----AAGTGAATGATCTCTACTGAGAAAGCTGC----- 784
Db 507 --ValCysAspSerValAsnAspCysGlyAspAsnSerAspGlyGlnGlyCysSerCysP 526
OY 785 -----CAGCAGCTGG 794
Db 526 roAlaGlnThrPheArgCysSerAsnGlyLysCysLeuSerLysSerGlnGlnCysAsnG 546
OY 795 GTTTCGAGAGTGTCTACCGGACAAACGAGTGTCCACACAGGAAATTTGCCAACAGCTTCT 854
Db 546 LysAspAspAspCysGlyAspGlySerAspGlyAlaSerCysProLysValAsnValT 566
OY 855 CAATCTGAGATCAACTCCACCATCCAGGAAAGCTTCCACAGCTCT-----G 902
Db 566 hrCysThrLysNIleThrTyrArgCysLeuAsnGlyLeuCysLeuSerLysGlyAsnProG 586
OY 903 AATGCCCTTCCACGCGGTATATCTCCCTCCAGCTGTCC----- 940
Db 586 lncysAspGlyLysGln-----AspCysSerAspGlySerAspGlyLysAspC 602
OY 941 --CACTGCGGATGAGGCGCATGAC-----GGGCGGATGCGGAGGCGGCTGCGCT 992
Db 602 ysaAspCysGlyLeuArgSerPheThrArgGlnAlaArgValAlaGlyGlyThrAspAla 622
OY 993 CGGATACCAAGTGGCTTTGGAAAGTGTGCAC--TTGGGACCAACCATCTGTG 1049
Db 622 spGlnGlyGlnTyrProTyrGlnValSerLeuNIleAlaLeuGlyGlnGlyNIleCysG 642
OY 1050 GAGGACGCTCATTTGACGCGGATGAGTGTGCTGACGCGGCGGCTGCTTCTGTCACCC 1109
Db 642 lylAspSerLeuIleSerProAsnTyrLeuValSerAlaNIleNIleCysTyrIleAspAsp 662
OY 1110 GGGAGAAAGTCTTGAGAGGC-----TGGAAGGTGTACGCGG 1145
Db 662 rg-----GlyPheArgTyrSerAspProThrGlnTyrThrAlaPheLeuG 677
OY 1146 GCACACCAACTGTCCACAGTGTGCTGAGGAGCTCC-----A 1184
Db 677 ly-----LeuNIleAspGlnSerGlnArgSerAlaProGlyValGlnGlnLysArgT 694
OY 1185 TTGGCGGATCATCATCAACCAATTAACACCAATGAGAGAGAGGATGATGACATGCGCC 1244
Db 694 euLysAlaGlyIleIleSerNIleProPhePheAsnAspPheThrPheAspTyrAlaIleAl 714
OY 1245 TCATGCGGCTGTCCAAAGCCCTGAACCTGTGCTGCTCATCATGCTGCTGCTGCCA 1304
Db 714 euLeuGlnLeuGlnLysProAlaGlnLysArgSerSerMetValArgProIleCysLeuPro 734
OY 1305 TGCAATGACAGACTTTTAGCTCATGAGACTCTGATGATACAGGCTTTGGCAGAGCA 1364
Db 734 spAlaSerNIleValPheProAlaGlyLysAlaIleTyrValThrGlyTyrPylNIleThr 753

```

```

OY 1365 GGGAGACATGACAAACATCCCTTCCGCGGAGGTGACAGTCAATCTCATGACT 1424
Db 754 --GlnTyrGlyGlyThrGlyAlaLeuIleLeuGlnLysGlyGlnIleArgValIleAsnG 773
OY 1425 TCAAGAAATGCAATGACTACTTGTGTTATGACACTTACCTTACCCCAAGATATGTGTG 1484
Db 773 lnrThrCysGlnAsnLeu-----ProGlnGlnIleThrProArgMetMetCysV 791
OY 1485 CTGGGAGCTTCCGCGGCGGACAGACTCTGCGGAGGAGACAGCGGGGGCTCTT---G 1541
Db 791 alGlyPheLeuSerGlyGlyValAspSerCysGlnGlyAspSerGlyGlyProLeuSer 811
OY 1542 TCTGTGACGAGAACACACGCTGTGATCCTGTGACAGGTGTACACAGCTGGGACAGCTGTG 1601
Db 811 erValGlnAlaAspGlyArgIlePheGlnAlaGlyValAlaSerTyrGlyAspGlyCysA 831
OY 1602 GCCAGAGAAACAAACCTGTGTGTGACACCAAGTGAAGAGAGTTCCTCGGATT 1657
Db 831 lagInArgAsnLysProGlyValTyrThrArgLeuProLeuPheArgAspTyrIle 849

RESULT 11
US-11-303-608-1
/ Sequence 1, Application US/11303608
/ Publication No. US20060171884A1
/ GENERAL INFORMATION:
/ APPLICANT: Folz, Ian
/ APPLICANT: King, Chadwick
/ APPLICANT: Ling, Peter Koon Bong
/ APPLICANT: Kang, Jaepal Singh
/ APPLICANT: Manchulenko, Kathy
/ APPLICANT: Chen, Francine
/ APPLICANT: Scatena, Caroline Darne
/ APPLICANT: Key, Bruce A.
/ APPLICANT: Madison, Edwin R.
/ APPLICANT: Godfrey, Wayne R.
/ APPLICANT: Morikowski, Stanislaw K.
/ APPLICANT: Richardson, Jennifer H.
/ TITLE OF INVENTION: BINDING PROTEINS SPECIFIC FOR HUMAN
/ TITLE OF INVENTION: MATRIXASE
/ FILE REFERENCE: ARGENTIX.148A
/ CURRENT APPLICATION NUMBER: US/11/303.608
/ PRIOR FILING DATE: 2005-12-16
/ PRIOR APPLICATION NUMBER: 60/706467
/ PRIOR FILING DATE: 2005-08-08
/ PRIOR APPLICATION NUMBER: 60/637859
/ PRIOR FILING DATE: 2004-12-20
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 855
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-303-608-1

Alignment Scores:
Pred. No.: 2,03e-24 Length: 855
Score: 532.50 Matches: 163
Percent Similarity: 40.0% Conservative: 61
Best Local Similarity: 29.1% Mismatches: 175
Query Match: 16.0% Indels: 161
DB: 7 Gaps: 25

US-10-806-370-11 (1-1748) x US-11-303-608-1 (1-855)
OY 282 TTTCAGGCTCTATCCGAGGTCATCATCCGACAGTACGCTCGGTACAACTGCC 341
Db 351 PheAsnSerProCysTyrTyrProGlyNIleTyr-----ProPro 362
OY 342 CAACCAAGTGTACTTGTATGAGCAACACCAAGTGGGGGCTGTACCATCCGATCATCTC 401
Db 363 AsnIleAspCysThr-----TyrAsnIleGlyValProAsnAsn--- 375
OY 402 CTGCGAGTGCAGC-----CAGCAACGAGGCGCA 431

```

```

Db      376 -----GlnHisValIysValArgPheIysPhePheThyLeuLeuGluProGlyVal 392
QY      432 CCAAGGAGAGCCCAAGTACAGAGCTGCCCAATTCA----- 467
Db      393 Pro-----AlaGlyThrCysProIysAspTyrValGluIleAsnGlyIuIys 408
QY      468 -----CCTGGCGGAGAGGCCAAG----- 485
Db      409 TyrCysGlyIuIysSerGlnPheValValThrSerAsnSerAsnIleThrValArg 428
QY      486 -----AGAGTACACCTCATCGGGTGCAGTCTCCCTCATTTGCTCCCTGGTGG 533
Db      429 PheHisSerAspGlnSerTyrThrAspThrGlyPheIuIaGluIuIyLeu----- 445
QY      534 TTTTGCATCATCTCTTCCAGTTCGTGGAGGGCCACAGAGGATGAGTACAGAGAGC 593
Db      446 -----SerTyrAspSerSerAspProCysProGlyGlnPheThrCysArg 460
QY      594 AGAGGAGAGCTGTCCCAAGACACCTGTGCTGTGACGGGTGTGAGTGTGCAAGCTGA 653
Db      461 ThrGlyArg--CysIleArgIysGluLeuArgCysAspGlyTTPAlaAspCysThrAspH 480
QY      654 AGAGTACAGAGCTGGGTGGGTGAGTTTGTACGTGGACAAGTCTGTGTTAAATCTACT 713
Db      480 IserAspGlyLeuAsnCys-----Serc 488
QY      714 CTGGGCTCTCCATCAG-----TGCGCTTC 737
Db      488 ysaAspAlaGlyHisIleGlnPheThrCysIysAsnIysPheCysIysPheLeuPheTrp----- 506
QY      738 CCATGTGTAGCAGC-----NACTGGAATGACTCTCTACTAGAGAAGACCTGC----- 784
Db      507 --ValCysAspSerValAsnAspCysGlyAspAsnSerAspGluGlnGlyCysSerCysP 526
QY      785 -----CAGAGCTGG 794
Db      526 roAlaGlnIthrPheArgCysSerAsnGlyIysCysLeuSerIysSerGlnIuIysAsnG 546
QY      795 GTTTCGAGAGTGTACCCGAGCAACGAGGTGGCCACAGGATTTTGGCAACAGCTTCT 854
Db      546 IlyAspAspAspCysGlyAspGlySerAspGluIaIasSerCysProIysValAsnValaI 566
QY      855 CAATCTTGAGATACACTCCACCATCCAGAAAGCTTCACAGCTCT-----G 902
Db      566 hrCysThrIysHisIleThrTyrArgCysLeuAsnGlyLeuCysLeuSerIysIysAsnProG 586
QY      903 AATGCCCTTCCACGCGATATCTCCCTCCAGTGTCC----- 940
Db      586 IuCyAspAspGlyIysGln-----AspCysSerAspGlySerAspGluIysAspC 602
QY      941 --CACTGCGAGTACAGGAGCCATGAC-----GGCGGATGTGGAGAGGGCGCTGGCT 992
Db      602 ysaAspCysGlyLeuArgSerPheThrArgGlnAlaArgValValGlyGlyThrAspAla 622
QY      993 CGGATACCAAGTGGCTTTGGCAAGTGAAGTCTGCAC--TTGGGACCAACCCACATCTGTG 1049
Db      622 spGluGlyGluTrpProItrpGlnValSerLeuHisAlaIleuGlyGlnGlyHisIleCysG 642
QY      1050 GAGGCAAGCTCATGAGCCGAGGAGGAGGCTCAGTGGCGCCACATGGTCTTCGTGACCC 1109
Db      642 IylAserIleuIleSerProAsnItrpLeuValIserAlaAlaHisCysTyrIleAspAsp 662
QY      1110 GGGAGAAAGTCTCTGAGAGGC-----TGGAAGGTGTACCGG 1145
Db      662 rg-----GlyPheArgTyrSerAspProThrGlnIthrPheAlaPheLeuG 677
QY      1146 GCACACCAAGCTGTGACAGTGTGCTGAGGACGCTCC-----A 1184
Db      677 Iy-----LeuHisAspGlnSerGlnArgSerAlaProGlyValGlnGluIuIyArgI 694
QY      1185 TTGCCGAGATCATGATCAACAGCAATTTCACCCGATGAGAGAGAGACTATGACATCGGCC 1244

```

```

Db      694 eulysArgIleIleSerHisProPhePheAsnAspPheThrPheAspTyrAspIleAla 714
QY      1245 TCATGCGGCTGTCCAAAGCCCTGACCCCTGTCCGCTACATCCACCTCTTGGCTCCCA 1304
Db      714 eulGluIleuGluIysProIaGluIuIyTrSerSerMetValArgProIleCysIeuProA 734
QY      1305 TGCATGAGACAGACTTTAGCTCAATGAGACTGTGCTGATACAGAGGCTTGGCAAGACA 1364
Db      734 spAlaSerHisValPheProAlaGlyIysValIleItrpValIthrGlyTrpGlyHisIthr 753
QY      1365 GGGAGACAGATGACAAAGACATCCCTCTCTCCGAGAGTGCAGAGTCAATCTCATGACT 1424
Db      754 --GlnTyrGlyIuIyThrGlyAlaIleuIleuGlnIysGlyIuIleArgValIleAsnG 773
QY      1425 TCAAGAAATGCATGACTACTGTCTATATACAGTTACCTTACCCCAAGATGATGTGTG 1484
Db      773 IthrThrCysGluAsnIleuLeu-----ProIuGlnIuIleThrProIArgMetMetCysv 791
QY      1485 CTGGGACCTTCTGGGGGCGAGAGACTCTGCCAGAGGACAGGGGGGCTCTT---G 1541
Db      791 alGlyPheIeuSerGlyIysValAspSerCysGlnIysAspSerGlyIysProIeuSers 811
QY      1542 TCTGTGACAGAAACAACCGCTGTGTACTGTGACAGGTGTACACAGCTGGGCAAGCTGTG 1601
Db      811 ervalGluIaAspGlyArgIlePheGlnIaGlyValIserTrpGlyAspGlyCysA 831
QY      1602 GCCAGAGAAACAACCTGTGTGTATCACCAAGTGCAGAGAAGTTCTTCCCTGATT 1657
Db      831 lagIuIyAsnIysProGlyValIyTrArgIeuProIeuPheArgAspTrpIle 849

RESULT 12
US-11-254-185-6
: Sequence 6, Application US/11254185
: Publication No. US20060099625A1
: GENERAL INFORMATION:
: APPLICANT: CRAIK, CHARLES S.
: APPLICANT: TAKEUCHI, TOSHIO
: APPLICANT: SCHUMAN, MARC
: TITLE OF INVENTION: MT-SPI SERINE PROTEASE
: FILE REFERENCE: 28644-701.302
: CURRENT APPLICATION NUMBER: US/11/254,185
: CURRENT FILING DATE: 2005-10-18
: NUMBER OF SEQ. ID NOS: 83
: SOFTWARE: PatentIn version 3.3
: SEQ. ID NO: 6
: LENGTH: 255
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:
: OTHER INFORMATION: Protein fragment/domain
US-11-254-185-6

Alignment Scores:
Pred. No.: 255
Score: 5,04e-24 Length: 255
Percent Similarity: 57.7% Matches: 105
Best Local Similarity: 42.3% Mismatches: 88
Query Match: 15.8% Indels: 17
Gaps: 6

US-10-806-370-11 (1-1748) x US-11-254-185-6 (1-255)
QY      971 ATGCTGGAGAGGGCGCTGCGCTCGATAGCAAGTGGCTTTGGCAAGTGTGCACTTC 1030
Db      1 IlevAlGlyIysArgAspTrpSerIeuGlyArgTrpProItrpGlnValIserIeuArgTyr 20
QY      1031 GGCACCAACCCACATCTGTGAGGAGCGCTCATTTAGCCGCCAGTGGGTGCTCACTGGCGCC 1090
Db      21 AspGlyAlaHisIeuCysGlyIysSerIeuIeuSerCysIysAspTrpValIeuIthrAlaIa 40
QY      1091 CACTGCTTCTTGTGACCCGCGAGAAAGCTCTGAGAGGCTGGAAGGTGTACGCGGACCC 1150
Db      41 HisCys---PheProIuIyArgAsnArgValIeuSerArgTrpArgValPheAlaGlyAla 59

```


Pred. No.:	5 04e-24	length:	25
Score:	525.50	Matches:	10
Percent Similarity:	57.7%	Conservative:	38
Best Local Similarity:	42.3%	Mismatches:	18
Query Match:	15.8%	Indels:	17
DB:	7	Gaps:	6

US-10-806-370-11 (1-1748) x US-11-253-869-6 (1-255)

OY	971	ATCTGGGAGGGGGCTGGCTGGCGATTTAGCAAGTGGCTTGGCAAGTAGTGTGCACTTC	103
Db	1	ITLeVtGtGtYtAgtAspThrSerLeuGtYtArgTrpProTrpGlnValSerLeuAgtGlyr	20
OY	1031	GGCAACCCACATCTGTGTGAGGAGCAGCTCATTTGACGCCCAAGTGGTCTCACTGCCGC	1090
Db	21	AspGtYtAItHtSLeuGtGtYtSerLeuSeruSerGtYtAspTrpValLeuThrAlaIa	40
OY	1091	CACGTCTTTCTTGTGTGACCCGGGAGAGTCTCTGAGGGCTGGAAAGTGTACCGGGCACC	1156
Db	41	HtScys---PheProGluAtrGAsnAgtValLeuSerAgtTrpArgValIleAlaGtYtA	59
OY	1151	-----ACCAACCTGCACAGTGTGCTGAGGAGCGCTCATCGCCAGATCATATCTC	1201
Db	60	ValIaGtInAlaSerProHtSgItyLeu-----GtLeuGtYtValGlnAlaValAItYr	77
OY	1202	AACGACCAATTAC-----ACCGATGAGGAGGACGACATATGACATCTGCC	1243
Db	78	HtSgItyGtYtYrLeuProPheAtrGAspTrpAsnSerGtGtInuAenSerAAsnIleA	97
OY	1244	CTCATGCGGCTGTCCAGCCCTCTGACCTGTGCCCTCAATCACTCACCTGTGGCTCCC	1303
Db	98	LeuValHtSLeuSerSerProLeuProLeuThrGtYrIleGlnProValCysLeuPro	117
OY	1304	ATGCATGAGCAGACCTTTAGCCTCAATGAGACCTGTGATGACAGGCTTTGGCAAGAC	1365
Db	118	AlaIaGtYtGlnAlaLeuValAspYtYtSLeuGtYtSerValHtGtYtRbYtAsnThr	137
OY	1364	AGGAGACAGATGACAAAGACATCCCTCTCTCCGGAGGTGCAAGTCAATCTCATGCAC	1423
Db	138	GltYrYtYrGtYtGlnGlnIaGtYt---ValLeuGtGlnIuAlaArgValProIleSer	156
OY	1424	TTCAAGAAATGCATGACTACTTGGTCTATGACAGTTACTTTACCCAGAGATGATGTGT	1483
Db	157	AsnAspValCysAsnGtYtAlaAspPheYtYrGtYtAsnGlnIleYtSProLYtSMeCHeCys	176
OY	1484	GCTGGGAGCCTTCGTGGGGGAGAGACCTCCGCGAGGAGGACAGCGGGGGGCTCTTGTCT	1543
Db	177	AlaGtYtYrProGtInuGtYtYtIleAspAlaCysGtInuGtYtAspSerGtYtYtProPheVal	196
OY	1544	TGTGTAG-----CAGAACAAACCGCTGTACTCTGGCAGGGGTCAACAGCTGGGGGC	1591
Db	197	CysGtInuAspSerIleSerAtrGtThrProAtrGtTrpArgLeuCysGtYtIleValSerTrpGtYt	216
OY	1592	ACAGGCTGTGGCCACAGAAACAAACCTGTGTGTACACCAAGTACAGAAATGTTCTTCC	1651
Db	217	ThrGtYtCysAlaLeuAlaGtInuYrProGtYtYtYrThrLYtSValSerAAspAgtGtInu	236
OY	1652	TGGATTTACACCAAGTAGAGAC 1675	
Db	237	TrpIlePheGlnAlaIleYtSThr 244	

RESULT 15
US-11-253-869-38
; Sequence 38, Application US/11253869
; Publication No. US20060104979A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; APPLICANT: SCHUMAN, MARC
; TITLE OF INVENTION: WT-SPL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 28644-701.303
; CURRENT APPLICATION NUMBER: US/11/253, 869
; CURRENT FILING DATE: 2005-10-18

```

; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatencIn version 3.3.3
; SEQ ID NO 38
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-253-869-38

```

Alignment Scores:

```
Pred. No.: 5.04e-24
Score: 525.50
Percent Similarity: 57.7%
Best Local Similarity: 42.3%
Query Match: 15.8%
DB: 7
```

US-10-806-370-11 (1-1748) x US-11-253-869-38 (1-255)

QY	971	ATGTGGAGAGGGGGCGTGGCGGTGGATAGAGAGGGCGTGGCAAGAGTCTCACTTC	1030
			:::
Db	1	lleva1elgylgylargabpnrtselueuglyargrtrprotrpnlvalserleuAlytr	20
			:::
QY	1031	GGCAACCCACACATCTGTGGAGGACGACCTCATTTAGACGCCACAGTGGTGTCTACTGCCGC	1090
			:::
Db	21	AspG1yAlaHnIsleuCGylgylserleuLeuSerGlyAspTrValleuthrAla1a	40
			:::
QY	1091	CACGTCTTTCTTGTGACCCGGGAGAAAGTCTGTGAGGGCTGGAAAGTGTACGGGGGACC	1150
			:::
Db	41	H1scYs---PhepProGluArGAsnaAvalleuSerArgrtrpArGvalPheAlaGlyAla	59
			:::
QY	1151	-----AGCAACCTGCACACAGTTGGCTGTGAGGACCTTCATGTGGCGAATCATATC	1201
			:::
Db	60	ValAlaGlnAlaSerProH1sGlyLeu-----GlnPheuglyAlaGlnAlaValAlaytr	77
			:::
QY	1202	AACGCAATATAC-----ACGATGAGGAGAGACGATATGAATATGGCC	1243
			:::
Db	78	H1sGlyGlyLyrrLeuProPheArGAspPProAsnSerGlnGluAsnSerAsnAsp1leAla	97
			:::
QY	1244	CTCATGGCGGTGTGTCCAGGCCCTGTGACCTGTGCGCTGCATCCACCTGTGCTGCCCC	1303
			:::
Db	98	LeuValH1sleuSerSerProLeuProLeuthrGlnTrGlnleuProVal1CyLeuPhePro	117
			:::
QY	1304	ATGATGAGACAGACCTTTAGCCTTCATAGAACACGTGTCGATACAGGCTTTGGCAAGACC	1363
			:::
Db	118	AlaAlaGlyGlnAlaLeuValAspGlyLyser1leCythrValHnrglyTrpGlyAsnThr	137
			:::
QY	1364	AGGAGAGACAGATGACAGACATCCCCCTCTCCGGAGGTGCAGGTCAATCTTCATGCAC	1423
			:::
Db	138	GlnTrYrrGlyGlnGlnAlaGly---ValleuGlnGlnAlaArgValPro1le1leSer	156
			:::
QY	1424	TTCAAGAAATGCATGACTACTGTGTTATGACAGTTAACTTACCCCAAGATATGTGT	1483
			:::
Db	157	AsnaAspVal1CyAsnGlyAla1aAspPheTrGlyAsnGln1le1yAsPro1ySmePheCys	176
			:::
QY	1484	GCTGGAGACCTCTGTGGGGGACAGACTCTGTGCACGAGAGACAGCGGGGGCTCTTGTC	1543
			:::
Db	177	AlaGlyLyrrProGlnGlnGly1leAspAla1CySelnGln1yAspSerGlyGlyProPheVal	196
			:::
QY	1544	TGTGAG-----GAGAACCCGCTGTGTACTGGCAGGTGTCCACGCTGGGGC	1591
			:::

```

OY      1592  ACAGGCTGTGGCCAGAGAACAAACCTGGTGTATACACCAAAGTACAGAAGTTCCTCC 1651
          |||||          |||||
DB      217  ThcIlycsalaleumalaglnlyspicgIlyValrYthrysvaIseraepheargln 236
          |||||          |||||
OY      1652  TGGATTTACAGCAAGATGAGAC 1675
          |||||          |||||
DB      237  TrpIlephclnalalrIelystnr 244
          |||||          |||||

```

Search completed: September 16, 2006, 02:48:37
Job time : 76 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:04:25 ; Search time 199 Seconds
(without alignments)
1291.234 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999
Sequence: 1 MERSHONASPARPSAGAS.....TEVLPWITYSKMSEVFRPKS 562

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_8: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *
9: Geneseqp2005s: *
10: Geneseqp2006s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	100.0	562	5	AAE17238 Human tra
2	2999	100.0	569	8	ADH17440 Human NOV
3	2991	99.7	562	4	AAE01943 Human tra
4	2991	99.7	562	7	AD110393 Human cel
5	2991	99.7	562	8	AD146917 Human tra
6	2991	99.7	562	9	ADY50146 Human end
7	2963	98.8	688	4	AAE01944 Human tra
8	2963	98.8	688	7	AD110395 Human cel
9	2963	98.8	688	8	AD146919 Human tra
10	2963	98.8	688	9	ADY50054 Human tra
11	2956	98.6	581	7	ADE31743 Human 291
12	2948.5	98.3	586	8	ADH17414 Human NOV
13	2948.5	98.3	586	8	ADH17434 Human NOV
14	2948.5	98.3	586	8	ADH17430 Human NOV
15	2943.5	98.1	586	8	ADH17452 Human NOV
16	2942.5	98.1	586	8	ADH17450 Human NOV
17	2942	98.1	581	5	ADZ5552 Human mos
18	2775.5	92.5	537	5	AAU82746 Amino aci
19	2775.5	92.5	537	8	ADH17438 Human NOV
20	2737.5	91.3	556	9	AAE20240 Novel hum
21	2686	89.6	542	8	ADH17432 Human NOV
22	2414.5	80.5	471	8	ABM8371 Human dia
23	2291.5	76.4	477	6	ABP56842 Human den

24	2287	76.3	421	4	AAB85042 Human SER
25	2287	76.3	421	8	ADH17442 Human NOV
26	2254	75.2	419	7	ADE29367 Human ser
27	2219.5	74.0	486	5	AAE19171 Human pro
28	2207	73.6	491	4	AAB93442 Human pro
29	2130.5	71.0	412	8	ADH17454 Human NOV
30	2108	70.3	446	6	ADP56843 Human den
31	2097	69.9	382	8	ADH17444 Human NOV
32	2059	68.7	406	8	ADH17436 Human NOV
33	2059	68.7	406	8	ADH17448 Human NOV
34	2026	67.6	401	8	ADH17422 Human NOV
35	2026	67.6	418	7	ADE79008 Human pro
36	2016	67.2	401	8	ADH17426 Human NOV
37	2013.5	67.1	449	7	ADE78981 Human pro
38	2012	67.1	401	8	ADH17428 Human pro
39	2008	67.0	451	7	ADE78982 Human pro
40	1906.5	63.6	375	7	ADE78993 Human pro
41	1842.5	61.4	370	8	ADH17424 Human NOV
42	1839.5	61.3	367	8	ADH17420 Human NOV
43	1311	43.7	296	3	AAV72108 Human ser
44	1311	43.7	372	3	AAV72092 Human ser
45	1311	43.7	372	5	ABG30789 Human ser

ALIGNMENTS

RESULT 1
ID AAE17238 standard; protein; 562 AA.
AC AAE17238;
XX
XX
DT 18-APR-2002 (first entry)
XX
XX
DE Human transmembrane serine protease.
XX
KW Human; transmembrane serine protease; gene therapy; metastasis; tumour;
KW chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;
KW atherosclerosis; neurodegenerative disease; neuroprotective; cyostatic;
KW pathogenic infection; antiinflammatory; antiarteriosclerotic;
KW antibacterial.
XX
XX Homo sapiens.
XX OS
XX WO200196538-A2.
XX
XX 20-DEC-2001.
XX
XX 12-JUN-2001; 2001WO-EP006618.
XX
XX 13-JUN-2000; 2000US-0211224P.
XX PR 13-APR-2001; 2001US-0283353P.
XX PR 16-APR-2001; 2001US-0283648P.
XX
XX (FARB) BAYER AG.
XX
XX Xiao Y, Gedrich R;
XX WPI; 2002-098065/13.
XX DR N-PSDB; AAD27734.
XX
XX Novel isolated polynucleotide encoding transmembrane serine protease
XX polypeptide, for treating chronic obstructive pulmonary disease, tumor
XX angiogenesis, inflammation, atherosclerosis and neurodegenerative
XX disease.
XX Claim 1; Fig 1; 120bp; English.
XX
XX The present invention relates to an isolated polynucleotide encoding a
XX transmembrane serine protease polypeptide. Transmembrane serine protease
XX gene is useful in gene therapy. The invention also relates to a
XX pharmaceutical composition which is useful for modulating the activity of

transmembrane serine protease in a disease, such as chronic obstructive pulmonary disease (COPD), metastasis of malignant cells, tumour angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease or pathogenic infection. Transmembrane serine protease is useful as a bait protein in a two-hybrid or three-hybrid assay. The polypeptide is useful for generating antibodies against it and in various assay systems. The present sequence is a human transmembrane serine protease

Sequence 562 AA;

Query Match 100.0%; Score 2999; DB 5; Length 562;

Best Local Similarity 100.0%; Pred. No. 8,2e-189; Mismatches 0; Indels 0; Gaps 0;

Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MERDGHNASPARTSAGASPAQASPAQTPPRASPAQASPAQASPAQTPPRASPAQAS 60
1 MERDGHNASPARTSAGASPAQASPAQTPPRASPAQASPAQASPAQTPPRASPAQAS 60
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAAPIRSSPARSAPATRAATESPSTSLPKFTREGGKQPLICCVLLALVYSLITL 180
121 PVGAAPIRSSPARSAPATRAATESPSTSLPKFTREGGKQPLICCVLLALVYSLITL 180
181 FQFMGHTGIRYKEQRESCPKIAVRCQGVDCCKLSDELGCVRFPMDSILKITYGSSHQ 240
181 FQFMGHTGIRYKEQRESCPKIAVRCQGVDCCKLSDELGCVRFPMDSILKITYGSSHQ 240
241 WLPICSSWNNDYSSEKTCQQLGFSFAHRTTEVAHRDPANFSILRYNSTIQSLHRSBCP 300
241 WLPICSSWNNDYSSEKTCQQLGFSFAHRTTEVAHRDPANFSILRYNSTIQSLHRSBCP 300
301 SORVSLSCSHCGEAMTGRIVGALASDSKMPQVSHFQTHICGTLIDAOVLTAA 360
301 SORVSLSCSHCGEAMTGRIVGALASDSKMPQVSHFQTHICGTLIDAOVLTAA 360
361 HCFVTRREKVLGKMYVAGTSMHLQPRASIAEIIINSNTYDEBDDVIALMRLSKPIAT 420
361 HCFVTRREKVLGKMYVAGTSMHLQPRASIAEIIINSNTYDEBDDVIALMRLSKPIAT 420
421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRTETDQKTSPLREVOVNLIDPKKNDYLV 480
421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRTETDQKTSPLREVOVNLIDPKKNDYLV 480
481 YDSYLTTPRMWCAQGLRGGRDSCQSGSPVCEQNNRWYLAGVTSWGTGCGGRNRPVYT 540
481 YDSYLTTPRMWCAQGLRGGRDSCQSGSPVCEQNNRWYLAGVTSWGTGCGGRNRPVYT 540
541 KYTEVLPMIYSMESEVFRFKS 562
541 KYTEVLPMIYSMESEVFRFKS 562

RESULT 2

ADH17440
ADH17440 standard; protein; 569 AA.

ADH17440;
11-MAR-2004 (first entry)

Human NOV12n protein - SEQ ID 130.

NOVX; antidiabetic; anorectic; cardiac; hypotensive;
antiartherosclerotic; anorectic; vinuclide; antibacterial; fungicide;
protoproctide; nootropic; neuroprotective; antiparkinsonian;
anticoronary; osteopathic; antiarthritic; antiinflammatory;
dermatological; antileukemic; antidiabetic; metabolic; diabetes;
obesity; infection; anorexia; cancer; cardiovascular; hypertension;
atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
epilepsy; immune; osteoarthritis; haemopoietic;

inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
cell differentiation; proliferation; haemopoiesis; wound healing;
angiogenesis; gene therapy; chromosome mapping; tissue typing;

pharmacogenomic; human.

Homo sapiens.

WO2003093432-A2.

13-NOV-2003.

02-MAY-2003; 2003WO-US013690.

02-MAY-2002; 2002US-0377321P.

08-MAY-2002; 2002US-0378730P.

24-MAY-2002; 2002US-0383075P.

29-MAY-2002; 2002US-0384044P.

30-MAY-2002; 2002US-0384215P.

30-MAY-2002; 2002US-0384296P.

30-MAY-2002; 2002US-0384327P.

30-MAY-2002; 2002US-0384352P.

31-MAY-2002; 2002US-0385211P.

02-JUL-2002; 2002US-0393333P.

09-AUG-2002; 2002US-0402154P.

09-AUG-2002; 2002US-0402171P.

09-AUG-2002; 2002US-0402204P.

22-AUG-2002; 2002US-0405175P.

27-AUG-2002; 2002US-0406125P.

23-SEP-2002; 2002US-0412954P.

30-SEP-2002; 2002US-0414975P.

07-OCT-2002; 2002US-0416661P.

24-OCT-2002; 2002US-0420851P.

31-OCT-2002; 2002US-0422547P.

01-MAY-2003; 2003US-00428275.

(CURA-) CURAGEN CORP.

Alvarez E, Anderson DM, Boldog FL, Catterton E, Edinger SR;

Fernandes ER, Gerlich VL, Gorman L, Grosse WM, Guo X, Ji W,

Kerkda R, Li L, Macdougall JR, Padigaru M, Patuwajan M,

Petererson JD, Raetelli L, Shinkets RA, Sytek KA, Stone DJ;

Vernet CAM, Voss EZ, Zhong M;

WPI; 2004-053040/05.

N-P8DB; ADH17439.

New isolated NOVX polypeptide, useful for preventing, diagnosing or

treating NOVX-associated disorders, e.g. osteoarthritis, obesity,

atherosclerosis, cancer, Parkinson's disease, asthma, or infections.

Claim 1; SEQ ID NO 130; 478bp; English.

The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates antidiabetic, anorectic, cardiac, hypotensive, antiatherosclerotic, anorectic, vinuclide, antibacterial, fungicide, protoprotide, nootropic, neuroprotective, antiparkinsonian, anticoronary, osteopathic, antiarthritic, antiinflammatory, dermatological, antileukemic and antidiabetic activities. The polypeptide of a medicament for treating metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders such as osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and various types of dyslipidaemia. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit neurogenesis, cell differentiation, cell proliferation, haemopoiesis, wound healing and angiogenesis, in gene therapy and the in generation of antibodies that bind immunospecifically to NOVX substances for use in therapeutic or

CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVA protein of the invention.

XX Sequence 569 AA:

Query Match 100.0%; Score 2999; DB 8; Length 569;
Best Local Similarity 100.0%; Pred. No. 8,3e-189;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQAS 60
DB 5 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQAS 64
QY 61 PACTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
DB 65 PACTPPGASPGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 124
QY 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGKOLPLIGCVLLIALVSLTLL 180
DB 125 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGKOLPLIGCVLLIALVSLTLL 184
QY 181 FQFMOGHTGIRYKQRESQPGHAYRCDGVNCKLSDELGCVRPDMDKSLIKYSGSSHQ 240
DB 185 FQFMOGHTGIRYKQRESQPGHAYRCDGVNCKLSDELGCVRPDMDKSLIKYSGSSHQ 244
QY 241 WLPICSSNMNDSYEKTCQQLGFESAHRTTEVAHRDFANSFSLIRYNSTIOESLHRSFCP 300
DB 245 WLPICSSNMNDSYEKTCQQLGFESAHRTTEVAHRDFANSFSLIRYNSTIOESLHRSFCP 304
QY 301 SQRYSISLQCSHCGIRAMTGRIVGALASDSKPMQVSLHFGTTHCGSTLIDAOVYLTAA 360
DB 305 SQRYSISLQCSHCGIRAMTGRIVGALASDSKPMQVSLHFGTTHCGSTLIDAOVYLTAA 364
QY 361 HCFVTRREKVLGEMKVYAGTSLNLHLPASIAEIIINSNTDEEDDIALMRLSKPLT 420
DB 365 HCFVTRREKVLGEMKVYAGTSLNLHLPASIAEIIINSNTDEEDDIALMRLSKPLT 424
QY 421 LSAHHPACLPAGHGTFSNETCWTGFGKTRTDDKTSPLRVEQVNLIDPKKNDYLV 480
DB 425 LSAHHPACLPAGHGTFSNETCWTGFGKTRTDDKTSPLRVEQVNLIDPKKNDYLV 484
QY 481 YDSVLTTPMMGAGDLRGGRDSCQSGSGPLVCEONNRVYLAGVMSWGCGCORMKRGVYT 540
DB 485 YDSVLTTPMMGAGDLRGGRDSCQSGSGPLVCEONNRVYLAGVMSWGCGCORMKRGVYT 544
QY 541 KYTEVLFWYSKMESEVFRKS 562
DB 545 KYTEVLFWYSKMESEVFRKS 566

RESULT 3
AAE01943 standard; protein; 562 AA.

XX AAE01943;

XX 31-JUL-2001 (first entry)

XX Human transmembrane serine protease (Endothelinase 2-S) protein.

XX Human, endothelinase 2-S; protease domain; cytosolic; vulnery; wound;
XX noctropic; peridontitis; dermatological disorder; gene therapy; scar;
XX angiogenesis; cardiovascular disorder; psoriasis; neovascular disease;
XX chronic inflammatory disease; ocular disorder; circulatory disorder;
XX crest syndrome; atherosclerosis; haemangiomatosis; diabetes mellitus;
XX liver cirrhosis; osteoradionecrosis; systemic sclerosis; oesophagial;
XX inflammatory bowel disease; fracture; rheumatoid arthritis; retinopathy;
XX systemic vasculitis; scleroderma; neoplasm; ulcer; burn;
XX transmembrane serine protease.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Domain 321..562
FT /label= Protease_domain

XX WO200136604-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000MO-US031803.

XX 18-NOV-1999; 99US-0166391P.

XX 22-SEP-2000; 2000US-0234840P.

XX (CORV-) CORVAS INT INC.

XX Madison EL, Ong EO;

XX WPI; 2001-336001/35.

XX N-PSDB; AAD05796.

XX New nucleic acid encoding a protein comprising endothelinase activity,
XX useful in the prevention and treatment of e.g. vascular malformations,
XX cardiovascular disorders, and chronic inflammatory disease.

XX Claim 4; Page 138-139; 152pp; English.

XX The present sequence is human short form transmembrane serine protease
XX (endothelinase 2-S) protein. Endothelinase 2 is a type-II membrane-type
XX serine protease which has a transmembrane domain at the N-terminus,
XX followed by a single low density lipoprotein-A receptor domain and a
XX single scavenger-receptor cysteine-rich domain. The C-terminus of the
XX endothelinase 2 contains the trypsin-like serine protease domain
XX characterised by the catalytic triad residues in 3 highly conserved
XX regions of the catalytic domain. In addition 3 repetitive sequence are
XX found just before the transmembrane domain and represent a sequence motif
XX for N-myristoylation modification. The invention relates to an
XX endothelinase protein, endothelinase protease domain and their
XX corresponding nucleic acid molecules. An endothelinase protein or protease
XX domain of it is useful for the treatment and diagnosis of disorders
XX associated with aberrant angiogenesis or undesired neovascularisation.
XX The undesired angiogenesis is associated with disorders selected from
XX solid neoplasia, vascular malformations and cardiovascular disorders such
XX as angiodioma, angiolioma, atherosclerosis, restenosis/reperfusion
XX injury, arteriovenous malformations, haemangiomatosis and vascular
XX adhesions, dyschondroplasia with vascular hamartomas (Fauci's
XX syndrome), hereditary haemorrhagic telangiectasia (Rendu-Osler-Weber
XX syndrome) and Von Hippel Lindau syndrome, chronic inflammatory bowel
XX disease, nonhealing fractures, peridontitis, psoriasis, rheumatoid
XX arthritis, venous stasis ulcers, granulations-burns, hypertrophic scars,
XX liver cirrhosis, osteoradionecrosis, postoperative adhesion, pyogenic
XX granuloma and systemic sclerosis and aberrant wound repairs, circulatory
XX disorders Raynaud's phenomenon, crest syndrome such as callosities,
XX oesophagial, dyomeolysis, sclerodactyly and telegiectasis, dermatological
XX disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
XX vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, port-wine
XX stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
XX caused by ocular neovascular disease, corneal graft neovascularisation,
XX macular degeneration, retinopathy of prematurity, retrolental fibroplasia
XX and corneal neovascularisation. The nucleic acids of the invention are
XX also used in gene therapy. The invention also provides method for
XX screening compounds that modulate angiogenesis

XX Sequence 562 AA:

Query Match 99.7%; Score 2991; DB 4; Length 562;
Best Local Similarity 99.8%; Pred. No. 2,8e-188;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MERDSHGASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQAS 60
|||||

```

Db 1 MERDSHGNSAPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Qy 61 PACTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
Db 61 PACTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
Qy 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLILALVLSLIL 180
Db 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLILALVLSLIL 180
Qy 181 FQFMOGHTGIRYKQORSCPHAVRCQDVVOCUKSDGLGCVRPMDKSLIKTYSGSSHQ 240
Db 181 FQFMOGHTGIRYKQORSCPHAVRCQDVVOCUKSDGLGCVRPMDKSLIKTYSGSSHQ 240
Qy 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSCEP 300
Qy 301 SQRVYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Db 301 SQRVYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
Qy 361 HCFVTRREKVLKLGKMTVAGTGNLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Db 361 HCFVTRREKVLKLGKMTVAGTGNLHQLPBAASIAEIIINSNTDEDDVDIALMLSKPLT 420
Qy 421 LSAHIHPACLPMMHGTQFSLNETCWTGFGKTRTDDTKTSPFLREYQVNLIDFKKCNLYLV 480
Db 421 LSAHIHPACLPMMHGTQFSLNETCWTGFGKTRTDDTKTSPFLREYQVNLIDFKKCNLYLV 480
Qy 481 YDSYLTTPRMWCAQDLRGGRDSCQDSDGGLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
Db 481 YDSYLTTPRMWCAQDLRGGRDSCQDSDGGLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
Qy 541 KYTEVLPMIYSKMSESVRFKRS 562
Db 541 KYTEVLPMIYSKMSESVRFKRS 562

RESULT 4
ADJ10393
ID ADJ10393 standard; proteain; 562 AA.
XX ADJ10393;
XX
XX 22-APR-2004 (first entry)
XX
XX Human cell surface protease #12.
XX
XX therapeutic agent; plasmin; protease specific antigen; PSA;
XX cell-surface protease-associated disease; cancer; ocular diseases;
XX cardiovascular disease; chronic inflammatory disease; wound;
XX circulatory disorder; dermatological disorder; rheumatoid arthritis;
XX psoriasis; diabetic retinopathy; pterygium;
XX excimer laser surgery scarring; glaucoma filtering surgery scarring;
XX macular degeneration; crest syndrome; solid neoplasm; vascular tumor;
XX melanoma; Kaposi's sarcoma; human; cell surface protease.
XX
XX Homo sapiens.
XX
XX WO200295007-A2.
XX
XX 28-NOV-2002.
XX
XX 23-MAY-2002; 2002WO-US016819.
XX
XX 23-MAY-2001; 2001US-0293267P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Slev DY,
XX
XX WPI; 2003-221280/21.

```

```

DR N-PSDB; ADJ10392.
XX
XX Novel conjugate useful for treating cell-surface protease-associated
XX disease, comprises a therapeutic agent and a peptidic or nucleic acid
XX substrate linked to it optionally by a peptidic linker.
XX
XX Claim 9; SEQ ID NO 24; 581bp; English.
XX
XX The invention comprises a conjugate that consists of a therapeutic agent
XX and a peptide substrate (optionally linked via linker). The peptide
XX substrate is proteolytically cleaved by a cell surface protease or a
XX soluble, released or shed form of it, to liberate the therapeutic agent,
XX the conjugate of the invention is not substantially cleaved by plasmin or
XX protease specific antigen (PSA). The conjugate of the invention is useful
XX for treating a cell-surface protease-associated disease such as: cancer,
XX ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
XX wounds, circulatory disorders, dermatological disorders, rheumatoid
XX arthritis, psoriasis, diabetic retinopathy, recurrence of pterygium,
XX scarring from excimer laser surgery, scarring from glaucoma filtering
XX surgery, macular degeneration, crest syndrome, solid neoplasm, vascular
XX tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence
XX represents a human cell surface protease.
XX
XX Sequence 562 AA:
XX
XX Query Match 99.7%; Score 2991; DB 7; Length 562;
XX Best Local Similarity 99.8%; Pred. No. 2.8e-188;
XX Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MERDSHGNSAPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
XX 1 MERDSHGNSAPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
XX 61 PACTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
XX 61 PACTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
XX 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLILALVLSLIL 180
XX 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLILALVLSLIL 180
XX 181 FQFMOGHTGIRYKQORSCPHAVRCQDVVOCUKSDGLGCVRPMDKSLIKTYSGSSHQ 240
XX 181 FQFMOGHTGIRYKQORSCPHAVRCQDVVOCUKSDGLGCVRPMDKSLIKTYSGSSHQ 240
XX 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSCEP 300
XX 241 WLPICSSNMNDYSSEKTCQOLGFESAHRTTEVAHRDPANSFSLRYNSTIOESLHRSCEP 300
XX 301 SQRVYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
XX 301 SQRVYISLQCSHCGLRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAA 360
XX 421 LSAHIHPACLPMMHGTQFSLNETCWTGFGKTRTDDTKTSPFLREYQVNLIDFKKCNLYLV 480
XX 421 LSAHIHPACLPMMHGTQFSLNETCWTGFGKTRTDDTKTSPFLREYQVNLIDFKKCNLYLV 480
XX 481 YDSYLTTPRMWCAQDLRGGRDSCQDSDGGLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
XX 481 YDSYLTTPRMWCAQDLRGGRDSCQDSDGGLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
XX 541 KYTEVLPMIYSKMSESVRFKRS 562
XX 541 KYTEVLPMIYSKMSESVRFKRS 562

RESULT 5
ADJ46917
ID ADJ46917 standard; proteain; 562 AA.

```


XX AC AD46917;
 XX DT 06-MAY-2004 (first entry)
 XX DE Human transmembrane serine protease (MTSP)-related polypeptide #2.
 XX
 KW Human; transmembrane serine protease; MTSP; cell surface protease;
 KW plasmin; prostate specific antigen; PSA; proliferative disease;
 KW cell-surface protease-associated disease; autoimmune disease;
 KW inflammatory disease; infectious disease; endocrine disease; cancer;
 KW ocular disorder; cardiovascular disorder; chronic inflammatory disease;
 KW wound; circulatory disorder; dermatological disorder; restenosis;
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy;
 KW laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; CREST syndrome; bacterial infection; viral disease;
 KW solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
 KW Kaposi's sarcoma; enzyme.
 XX
 OS Homo sapiens.
 XX PN US2004001801-A1.
 XX PD 01-JAN-2004.
 XX PE 23-MAY-2002; 2002US-00156214.
 XX PR 23-MAY-2002; 2002US-00156214.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Siev DV;
 DR WPI; 2004-190126/18.
 DR N-PSDB; ADJ46916.
 XX PT Conjugate useful for treating e.g. cancer, cell-surface protease-
 PT associated diseases, comprising a peptidic substrate or nucleic acid
 PT substrate linked to a therapeutic agent through a linker.
 PS Claim 9; SEQ ID NO 24; 361pp; English.
 XX
 CC The invention relates to a conjugate comprising a therapeutic agent and a
 CC peptidic substrate or nucleic acid substrate linked to the agent
 CC optionally through a linker or peptidic linker, where the peptidic
 CC substrate is proteolytically cleaved by a cell surface protease or a
 CC soluble, released or shed form conjugate to liberate the agent and the
 CC conjugate is not substantially cleaved by plasmin or prostate specific
 CC antigen (PSA). The conjugate is useful for treating a disease, which
 CC involves administering a conjugate to a subject, where the disease is
 CC preferably a proliferative disease or a cell-surface protease-associated
 CC disease. The diseases include autoimmune diseases, inflammatory diseases,
 CC infectious diseases and endocrine diseases. The conjugate is useful for
 CC treating a cell-surface protease-associated disease, which involves
 CC administering a conjugate comprising an agent and a peptidic substrate to
 CC a subject exhibiting symptoms of a cell-surface protease-associated
 CC disorder, where the disease is selected from cancer, ocular disorders,
 CC cardiovascular disorders, chronic inflammatory diseases, wounds,
 CC circulatory disorders, dermatological disorders, restenosis, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, scarring from laser
 CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
 CC CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
 CC vascular tumours such as lung cancer, colon cancer, prostate cancer,
 CC melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
 CC serine protease (MTSP)-related polypeptide of the invention.
 XX
 SQ Sequence 562 AA;
 Query Match 99.7%; Score 2991; DB 8; Length 562;
 Best Local Similarity 99.8%; Pred. No. 2.8e-188;
 Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MERDGHNASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 QY 61 PACTPPGRASPGRASPAPQASPARASPLASISRSRSGSSRSASASVTTSTRTYLVNAT 120
 DB 61 PACTPPGRASPGRASPAPQASPARASPLASISRSRSGSSRSASASVTTSTRTYLVNAT 120
 QY 121 PVGAVPIRSSPARSPAPTRATRESPGTSLPKFTREGOKOLPLGCVLLTALVSLIIL 180
 DB 121 PVGAVPIRSSPARSPAPTRATRESPGTSLPKFTREGOKOLPLGCVLLTALVSLIIL 180
 QY 181 FQFMQHTGIRKRYKQRESCPGHAIVRCGVVDCUKLSDELGCVRPDMDKSLUKIYSGSSHQ 240
 DB 181 FQFMQHTGIRKRYKQRESCPGHAIVRCGVVDCUKLSDELGCVRPDMDKSLUKIYSGSSHQ 240
 QY 241 WLPICSSNMWDSYEKTCQQLGFESAHRTTEVAHRDPANFSIIRYNSTIOESLHRSCEP 300
 DB 241 WLPICSSNMWDSYEKTCQQLGFESAHRTTEVAHRDPANFSIIRYNSTIOESLHRSCEP 300
 QY 301 SORYSISLOCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
 DB 301 SORYSISLOCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWVLTAA 360
 QY 361 HCFVTRKYLEGKVVYAGTSLNLHQLPEASIAEIIINSNTDEBDYDIALMRLSKPUL 420
 DB 361 HCFVTRKYLEGKVVYAGTSLNLHQLPEASIAEIIINSNTDEBDYDIALMRLSKPUL 420
 QY 421 LSAHIHPACLBMGQTSLSNETCWTITGFKTRFTDDKTSPLRLVQVNLIFPKKNDVLV 480
 DB 421 LSAHIHPACLBMGQTSLSNETCWTITGFKTRFTDDKTSPLRLVQVNLIFPKKNDVLV 480
 QY 481 YDSVLTPEPMKAGLRGGRDSCGDSGGLVCEQNNRYLAGVTSWGCGRGGRNKPQVYT 540
 DB 481 YDSVLTPEPMKAGLRGGRDSCGDSGGLVCEQNNRYLAGVTSWGCGRGGRNKPQVYT 540
 QY 541 KYTEVLPMIYSKMESEVFRKS 562
 DB 541 KYTEVLPMIYSKMESEVFRKS 562
 RESULT 6
 ID ADY50146 standard; protein, 562 AA.
 AC ADY50146;
 DT 19-MAY-2005 (first entry)
 DE Human endothelialase-2L.
 KW Antiarthritic; Antirheumatic; Antipsoriatic; Antidiabetic;
 KW Ophthalmological; Cardiovascular-Gen.; Vulnary; Antiinflammatory;
 KW Vasootropic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;
 KW rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
 KW dermatological disease; cancer; neoplasm; endothelialase-2; enzyme.
 OS Homo sapiens.
 PN WO2005019270-A2.
 PD 03-MAR-2005.
 PF 12-AUG-2004; 2004WO-US026148.
 PR 14-AUG-2003; 2003US-0495005P.
 PR 14-NOV-2003; 2003US-0520164P.
 PA (DYAX-) DYAX CORP.
 PA (DEND-) DENDREON CORP.
 PI Madison EL, Nixon A;
 DR WPI; 2005-202609/21.

DR N-PSDB; ADY50145.

PT Novel protein capable of inhibiting endothelinase-2, useful for treating
PT or preventing angiogenesis related disorder e.g. cancer.

PS Disclosure; SEQ ID NO 94; 157pp; English.

Disclosure; SEQ ID NO 94; 157pp; English.

The invention relates to an isolated protein (I) which comprises a heavy chain (HC) immunoglobulin variable domain sequence and a light chain (LC) immunoglobulin variable domain sequence, where the first and second immunoglobulin variable domain sequences form an antigen binding site that specifically binds to human endothelial-2. (I) is useful for detecting an endothelial or endothelial activity in a sample, for modulating an activity of an ER2-expressing cell, for modulating proteolysis, for killing or inhibiting growth of a cell, for detecting endothelial in a subject, for modulating endothelial activity in a subject, for treating or preventing a disorder characterized by unwanted angiogenesis in a subject. The disorder is chosen from rheumatoid arthritis, psoriasis, diabetic retinopathies, ocular disorder such as pterygial recurrence, scarring excimer laser surgery and glaucoma filtering surgery, cardiovascular disorders, chronic inflammatory disorders, wound repair, circulatory disorders, crest syndromes, dermatological disorders and cancers. The present sequence represents the human endothelial-2S.

The inventors relates to an isolated protein (I) which comprises a heavy chain (HC) immunoglobulin variable domain sequence and a light chain (LC) immunoglobulin variable domain sequence, where the first and second immunoglobulin variable domain sequences from an antigen binding site that specifically binds to human endothelase-2. (I) is useful for detecting an endothelase or endothelase activity in a sample, for modulating an activity of an ET-2-expressing cell, for modulating procoagulate, for killing or inhibiting growth of a cell, for detecting endothelase in a subject, for modulating endothelase activity in a subject, for treating or preventing a disorder characterized by unwanted angiogenesis in a subject. The disorder is chosen from rheumatoid arthritis, psoriasis, diabetic retinopathies, ocular disorder such as filerigil recurrence, scarring eximer laser surgery and glaucoma filtering surgery, cardiovascular disorders, chronic inflammatory disorders, wound repair, circulatory disorders, crest syndrome, dermatological disorders and cancers. The present sequence represents the human endothelase-2S.

CC disorders Raynaud's phenomenon, crest syndromes such as calcinosis,
 CC oesophageal, dyonocytosis, sclerodactyly and teanaglectasis, dermatological
 CC disorders such as systemic vasculitis, scleroderma, pyoderma gangrenosum,
 CC vasculopathy, venous, arterial ulcers, Sturge-Weber syndrome, Port-wine
 CC stains, blue rubber bleb nevus syndrome, Klippel-Trenaunay-Weber syndrome
 CC and Osler-Weber-Rendu syndrome and ocular disorders such as blindness
 CC caused by ocular neovascular disease, corneal graft neovascularisation,
 CC macular degeneration, retinopathy of prematurity, retrolental fibroplasia
 CC and corneal neovascularisation. The nucleic acids of the invention are
 CC also used in gene therapy. The invention also provides method for
 CC screening compounds that modulate angiogenesis

XX Sequence 688 AA:

Query Match 98.8%; Score 2963; DB 4; Length 688;
 Best Local Similarity 99.3%; Pred. No. 2.4e-186;

Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNSAPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 DB 1 MERDSHGNSAPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 QY 61 PACTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 DB 61 PACTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 QY 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 DB 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 QY 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 DB 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 QY 181 FQFMQGHGIRYKREKRESCPKHVAACDGVNDCKLSDELGCVRFPMDSLLKIYSGSSHQ 240
 DB 181 FQFMQGHGIRYKREKRESCPKHVAACDGVNDCKLSDELGCVRFPMDSLLKIYSGSSHQ 240
 QY 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSRCP 300
 DB 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSRCP 300
 QY 301 SORITSLQCHSCGIRAMTGRIVGALASDSKMPQOVSLHFGTTHICGGTLLIDAQWVLTAA 360
 DB 301 SORITSLQCHSCGIRAMTGRIVGALASDSKMPQOVSLHFGTTHICGGTLLIDAQWVLTAA 360
 QY 361 HCFPVTRKEVLEGMKVYAGTSLNHLQLEPAASIAEIIINSNTYDEEDVDIALMRLSKPLT 420
 DB 361 HCFPVTRKEVLEGMKVYAGTSLNHLQLEPAASIAEIIINSNTYDEEDVDIALMRLSKPLT 420
 QY 421 LSAHIHPACLPWAGTFSINETCWTGFGKTRRETDDKTSPLREVOVNLIDPKKCNLYLV 480
 DB 421 LSAHIHPACLPWAGTFSINETCWTGFGKTRRETDDKTSPLREVOVNLIDPKKCNLYLV 480
 QY 481 YDSYITPRMCAAGDLRGGRDSQQDSSGGLVCEQNNRYTLAGVTSWGTGCCGRNKKPGYTT 540
 DB 481 YDSYITPRMCAAGDLRGGRDSQQDSSGGLVCEQNNRYTLAGVTSWGTGCCGRNKKPGYTT 540
 QY 541 KYTEVLPMIYSKMESEVR 558
 DB 541 KYTEVLPMIYSKMESEVR 558

RESULT 8

AD110395 AD110395 standard; protein; 688 AA.

XX AC AD110395;

XX DT 22-APR-2004 (first entry)

XX DE Human cell surface protease #13.

XX therapeutic agent; plasmin; protease specific antigen; PSA;
 KW cell-surface protease-associated disease; cancer; ocular disease;
 KW cardiovascular disease; chronic inflammatory disease; wound;
 KW circulatory disorder; dermatological disorder; rheumatoid arthritis;

KW psoriasis; diabetic retinopathy; pterygium;
 KW excimer laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; crest syndrome; solid neoplasm; vascular tumour;
 KW melanoma; Kaposi's sarcoma; human; cell surface protease.

OS Homo sapiens.

PN WO200295007-A2.

XX 28-NOV-2002.

XX 23-MAY-2002; 2002WO-US016819.

XX 23-MAY-2001; 2001US-0293267P.

XX (CORV-) CORVAS INT INC.

XX Madison EL, Sempke JE, Vlasuk GP, Kemp SJ, Komandla M, Stev DV;

XX WPI; 2003-221280/21.

XX N-PSDB; AD110394.

PT Novel conjugate useful for treating cell-surface protease-associated
 PT disease, comprises a therapeutic agent and a peptidic or nucleic acid
 PT substrate linked to it optionally by a peptidic linker.

XX Claim 9; SEQ ID NO 26; 581bp; English.

CC The invention comprises a conjugate that consists of a therapeutic agent
 CC and a peptide substrate (optionally linked via linker). The peptide
 CC substrate is proteolytically cleaved by a cell surface protease or a
 CC soluble, released or shed form of it, to liberate the therapeutic agent,
 CC the conjugate of the invention is not substantially cleaved by plasmin or
 CC protease specific antigen (PSA). The conjugate of the invention is useful
 CC for treating a cell-surface protease-associated disease such as: cancer,
 CC ocular diseases, cardiovascular diseases, chronic inflammatory diseases,
 CC wounds, circulatory disorders, dermatological disorders, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, recurrence of pterygium,
 CC scarring from excimer laser surgery, scarring from glaucoma filtering
 CC surgery, macular degeneration, crest syndromes, solid neoplasms, vascular
 CC tumours, melanoma and Kaposi's sarcoma. The present amino acid sequence
 CC represents a human cell surface protease.

XX Sequence 688 AA:

Query Match 98.8%; Score 2963; DB 7; Length 688;
 Best Local Similarity 99.3%; Pred. No. 2.4e-186;

Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNSAPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 DB 1 MERDSHGNSAPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 QY 61 PACTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 DB 61 PACTPPGASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 QY 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 DB 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 QY 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 DB 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVLSLITL 180
 QY 181 FQFMQGHGIRYKREKRESCPKHVAACDGVNDCKLSDELGCVRFPMDSLLKIYSGSSHQ 240
 DB 181 FQFMQGHGIRYKREKRESCPKHVAACDGVNDCKLSDELGCVRFPMDSLLKIYSGSSHQ 240
 QY 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSRCP 300
 DB 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHDFANSFSLRYNSTIOESLHRSRCP 300
 QY 301 SORITSLQCHSCGIRAMTGRIVGALASDSKMPQOVSLHFGTTHICGGTLLIDAQWVLTAA 360
 DB 301 SORITSLQCHSCGIRAMTGRIVGALASDSKMPQOVSLHFGTTHICGGTLLIDAQWVLTAA 360

QY 361 HCFVTRKVLKGVKAVYAGTSLHQLPRAASIAIIINSNTDEDDYDIALMRLSKPLT 420
 Db 361 HCFVTRKVLKGVKAVYAGTSLHQLPRAASIAIIINSNTDEDDYDIALMRLSKPLT 420
 QY 421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRRETDKTSPLREVQVNLIDFKKNDVLY 480
 Db 421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRRETDKTSPLREVQVNLIDFKKNDVLY 480
 QY 481 YDSYLTTPPMKAGDLRGGRDSCQDGGPPLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
 Db 481 YDSYLTTPPMKAGDLRGGRDSCQDGGPPLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
 QY 541 KYTEVLPMWYSKMESEVR 558
 Db 541 KYTEVLPMWYSKMESEVR 558
 Db 541 KYTEVLPMWYSKMESEVR 558
 RESULT 9
 ADJ46919
 ID ADJ46919 standard; protein; 688 AA.
 AC ADJ46919;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human transmembrane serine protease (MTSP)-related polypeptide #3.
 XX
 KW Human; transmembrane serine protease; MTSP; cell surface protease;
 KW plasmin; prostate specific antigen; PSA; proliferative disease;
 KW cell-surface protease-associated disease; autoimmune disease;
 KW inflammatory disease; infectious disease; endocrine disease; cancer;
 KW ocular disorder; cardiovascular disorder; chronic inflammatory disease;
 KW wound; circulatory disorder; dermatological disorder; resectomias;
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy;
 KW laser surgery scarring; glaucoma filtering surgery scarring;
 KW macular degeneration; CREST syndrome; bacterial infection; viral disease;
 KW solid neoplasm; vascular tumour; lung; colon; prostate; melanoma;
 KW Kaposi's sarcoma; enzyme.
 XX
 XX Homo sapiens.
 OS
 PN US2004001801-A1.
 XX
 PD 01-JAN-2004.
 XX
 PF 23-MAY-2002; 2002US-00156214.
 XX
 PR 23-MAY-2002; 2002US-00156214.
 XX
 PA (CORV-) CORVAS INT INC.
 PI Madison EL, Sempke JB, Vlasuk GP, Kemp SJ, Komandla M, Siew DV;
 XX
 DR WPI; 2004-190126/18.
 DR N-PSDB; ADJ46918.
 XX
 PT Conjugate useful for treating e.g. cancer, cell-surface protease-
 PT associated diseases, comprising a peptidic substrate or nucleic acid
 PT substrate linked to a therapeutic agent through a linker.
 XX
 PS Claim 9; SEQ ID NO 26; 361pp; English.
 XX
 CC The invention relates to a conjugate comprising a therapeutic agent and a
 CC peptidic substrate or nucleic acid substrate linked to the agent
 CC optionally through a linker or peptidic linker, where the peptidic
 CC substrate is proteolytically cleaved by a cell surface protease or a
 CC soluble, released or shed form conjugate to liberate the agent and the
 CC conjugate is not substantially cleaved by plasmin or prostate specific
 CC antigen (PSA). The conjugate is useful for treating a disease, which
 CC involves administering a conjugate to a subject, where the disease is
 CC preferably a proliferative disease or a cell-surface protease-associated
 CC disease. The diseases include autoimmune diseases, inflammatory diseases,
 CC infectious diseases and endocrine diseases. The conjugate is useful for

CC treating a cell-surface protease-associated disease, which involves
 CC administering a conjugate comprising an agent and a peptidic substrate to
 CC a subject exhibiting symptoms of a cell-surface protease-associated
 CC disorder, where the disease is selected from cancer, ocular disorders,
 CC cardiovascular disorders, chronic inflammatory diseases, wounds, rheumatoid
 CC arthritis, psoriasis, diabetic retinopathy, scarring from laser
 CC surgery, scarring from glaucoma filtering surgery, macular degeneration,
 CC CREST syndrome, bacterial infections, viral diseases, solid neoplasms and
 CC vascular tumours such as lung cancer, colon cancer, prostate cancer,
 CC melanoma and Kaposi's sarcoma. This sequence represents a transmembrane
 CC serine protease (MTSP)-related polypeptide of the invention.
 XX
 XX Sequence 688 AA;
 SQ
 Query Match 98.8%; Score 2963; DB 8; Length 688;
 Best Local Similarity 99.3%; Pred. No. 2,4e-186;
 Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQTPPGRASPAQAS 60
 Db 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQTPPGRASPAQAS 60
 QY 61 PAGTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 Db 61 PAGTPPGRASPGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
 QY 121 PVGAVPISSPARAPATRAPRESPTSLPKFTREGQKOLPLIGCVLLIALVLSLIL 180
 Db 121 PVGAVPISSPARAPATRAPRESPTSLPKFTREGQKOLPLIGCVLLIALVLSLIL 180
 QY 181 FQFMQGHGIRYKQRESCEPHAVRCQGVNDCKLSDELGCVRPDMDSLKITYGSSHQ 240
 Db 181 FQFMQGHGIRYKQRESCEPHAVRCQGVNDCKLSDELGCVRPDMDSLKITYGSSHQ 240
 QY 241 WLPICSSNMNDSYSEKTCQOAGFESSAHTTEVAHDFPANSSTILRYNSTIGESLHRSCEP 300
 Db 241 WLPICSSNMNDSYSEKTCQOAGFESSAHTTEVAHDFPANSSTILRYNSTIGESLHRSCEP 300
 QY 301 SQRYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQVLTAA 360
 Db 301 SQRYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQVLTAA 360
 QY 361 HCFVTRKVLKGVKAVYAGTSLHQLPRAASIAIIINSNTDEDDYDIALMRLSKPLT 420
 Db 361 HCFVTRKVLKGVKAVYAGTSLHQLPRAASIAIIINSNTDEDDYDIALMRLSKPLT 420
 QY 421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRRETDKTSPLREVQVNLIDFKKNDVLY 480
 Db 421 LSAHHPACLPMHGQTFSLNETCWTGFGKTRRETDKTSPLREVQVNLIDFKKNDVLY 480
 QY 481 YDSYLTTPPMKAGDLRGGRDSCQDGGPPLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
 Db 481 YDSYLTTPPMKAGDLRGGRDSCQDGGPPLVCEQNNRWYLAGVTSWGTCCGQRNKPQVYT 540
 QY 541 KYTEVLPMWYSKMESEVR 558
 Db 541 KYTEVLPMWYSKMESEVR 558
 RESULT 10
 ADJ50054
 ID ADJ50054 standard; protein; 688 AA.
 AC ADJ50054;
 XX
 DT 19-MAY-2005 (first entry)
 XX
 DE Human endotheliasis-2L.
 XX
 KW Antiarrhythmic; Antirheumatic; Antipruritic; Antidiabetic;
 KW Ophthalmological; Cardiovascular-Gen.; Vulnerary; Antiinflammatory;
 KW Vasotrophic; Dermatological; Cytostatic; Antiangiogenic; angiogenesis;

KW rheumatoid arthritis; psoriasis; cardiovascular disease; inflammation;
 KW dermatological disease; cancer; neoplasm; endothelase-2; enzyme.
 OS Homo sapiens.
 XX WO2005019270-A2.
 XX PD 03-MAR-2005.
 XX PF 12-AUG-2004; 2004WO-US026148.
 XX PR 14-AUG-2003; 2003US-0495005P.
 XX PR 14-NOV-2003; 2003US-0520164P.
 XX PA (DYAX-) DYAX CORP.
 XX PA (DEND-) DENDREON CORP.
 XX PI Madison EL, Nixon A;
 XX MP1; 2005-202609/21.
 XX DR N-PSDB; ADY50053.
 XX PT Novel protein capable of inhibiting endothelase-2, useful for treating
 XX PT or preventing angiogenesis related disorder e.g. cancer.
 XX PS Disclosure; SEQ ID NO 2; 157pp; English.
 CC The invention relates to an isolated protein (1) which comprises a heavy
 CC chain (HC) immunoglobulin variable domain sequence and a light chain (LC)
 CC immunoglobulin variable domain sequence, where the first and second
 CC immunoglobulin variable domain sequences from an antigen binding site
 CC that specifically binds to human endothelase-2. (1) is useful for
 CC detecting an endothelase or endothelase activity in a sample, for
 CC modulating an activity of an E12-expressing cell, for modulating
 CC proteolysis, for killing or inhibiting growth of a cell, for detecting
 CC endothelase in a subject, for modulating endothelase activity in a
 CC subject, for treating or preventing a disorder characterized by unwanted
 CC angiogenesis in a subject. The disorder is chosen from rheumatoid
 CC arthritis, psoriasis, diabetic retinopathies, ocular disorder such as
 CC pterygii recurrence, scarring excimer laser surgery and glaucoma
 CC filtering surgery, cardiovascular disorders, chronic inflammatory
 CC disorders, wound repair, circulatory disorders, crest syndromes,
 CC dermatological disorders and cancers. The present sequence represents the
 CC human endothelase-2L.
 XX SQ Sequence 688 AA;
 Query Match 98.8%; Score 2963; DB 9; Length 688;
 Best Local Similarity 99.3%; Pred. No. 2.4e-186;
 Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 DB 1 MERDSHGNSAPARTSPASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
 QY 61 PAGTPPGASPRRASPAPASPAASLASRSSSSGSSSKRSASVTTTSPRYLVNAT 120
 DB 61 PAGTPPGASPRRASPAPASPAASLASRSSSSGSSSKRSASVTTTSPRYLVNAT 120
 QY 121 PVGAVPIRSSPARSPATRTATRESPTSLPKFTWREGKQPLICVLLILALVSLIIL 180
 DB 121 PVGAVPIRSSPARSPATRTATRESPTSLPKFTWREGKQPLICVLLILALVSLIIL 180
 QY 181 FQFWGHTGIRYKEQRESCPKIAVRCQGVNCKLSDELGCVRFPMWDXSLKITYGSSHQ 240
 DB 181 FQFWGHTGIRYKEQRESCPKIAVRCQGVNCKLSDELGCVRFPMWDXSLKITYGSSHQ 240
 QY 241 WPIPTSSNMNDSYSEKTCQQLGFESAHRTTEVAHDPANFSIILYNSTIOESLRSECP 300
 DB 241 WPIPTSSNMNDSYSEKTCQQLGFESAHRTTEVAHDPANFSIILYNSTIOESLRSECP 300
 QY 301 SQRYSILOCSSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAA 360
 DB 301 SQRYSILOCSSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAA 360

DB 301 SQRYSILOCSSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAA 360
 QY 361 HCFVTRPKVLEGMKVVAAGTNSLHQLPEAASIAIINSNTDEEDVDIALMRISKPLT 420
 DB 361 HCFVTRPKVLEGMKVVAAGTNSLHQLPEAASIAIINSNTDEEDVDIALMRISKPLT 420
 QY 421 LSAHIHPACLPMHQTFSLNETCWIITGFKRTRETDKTSPELREOVNLIIDFKKNDYLV 480
 DB 421 LSAHIHPACLPMHQTFSLNETCWIITGFKRTRETDKTSPELREOVNLIIDFKKNDYLV 480
 QY 481 YDSYLTFRMCMACGLRGGRDSCQDSSGAPLVCEONNRYLAGVTSWGTGCCORNKPGVYT 540
 DB 481 YDSYLTFRMCMACGLRGGRDSCQDSSGAPLVCEONNRYLAGVTSWGTGCCORNKPGVYT 540
 QY 541 KTEVLPMWYSKMESEVR 558
 DB 541 KTEVLPMWYSKMESEVR 558
 RESULT 11
 ID ADE31743 standard; protein, 581 AA.
 AC ADE31743;
 XX 29-JAN-2004 (first entry)
 DE Human 2914 protein #SEQ ID 100.
 XX Antiartherosclerotic; cardiant; vasotropic; antiinflammatory;
 KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
 KW cardiovascular; disorder; ischaemia; aortic bending;
 KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
 KW angina; cardiomyopathy; cardiac death.
 OS Homo sapiens.
 XX WO2003065984-A2.
 PD 14-AUG-2003.
 XX 29-JAN-2003; 2003WO-US002571.
 PR 01-FEB-2002; 2002US-0353224P.
 PR 15-MAR-2002; 2002US-0364529P.
 PR 19-APR-2002; 2002US-0373861P.
 PR 29-APR-2002; 2002US-0376287P.
 PR 12-JUN-2002; 2002US-0388080P.
 PR 24-JUN-2002; 2002US-0390971P.
 PR 03-JUL-2002; 2002US-0394130P.
 PR 10-JUL-2002; 2002US-0394797P.
 PR 21-AUG-2002; 2002US-0404904P.
 PR 23-AUG-2002; 2002US-0405450P.
 PR 04-SEP-2002; 2002US-0408070P.
 PR 06-NOV-2002; 2002US-0424300P.
 PR 05-DEC-2002; 2002US-0431042P.
 PR 05-DEC-2002; 2002US-0431079P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Logan TV, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
 PI Stigliano N, Perodin J, Rodrigue-Way A;
 DR WPI; 2003-731468/69.
 DR N-PSDB; ADE31742.
 XX Identifying a compound capable of treating a cardiovascular disorder
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193
 PT polypeptide or nucleic acid.
 PS Disclosure; SEQ ID NO 100; 328pp; English.
 XX

XX The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic, anorectic,
 CC cardiant, hypotensive, antiarteriosclerotic, anorectic, virucide,
 CC antibacterial, fungicide, protozoacide, neuroprotective, antiinflammatory,
 CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritic,
 CC antiinflammatory, dermatological, antiaesthetic and antipneumonic
 CC activities. The polypeptide, nucleic acid molecules and antibodies may
 CC be useful in the manufacture of a medicament for treating metabolic
 CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
 CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
 CC diseases including hypertension and atherosclerosis, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
 CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
 CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
 CC acids and polypeptides may also be used as targets for the identification
 CC of small molecules that modulate or inhibit neurogenesis, cell
 CC differentiation, cell proliferation, haemopoiesis, wound healing and
 CC angiogenesis, in gene therapy and the in generation of antibodies that
 CC bind immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids may be further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine and pharmacogenomics. The current sequence is that of the human
 CC NOVX protein of the invention.

XX
 SO Sequence 586 AA:

Query Match 98.3%; Score 2948.5; DB 8; Length 586;
 Best Local Similarity 99.1%; Pred. No. 1.8e-185;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MERDSHGASAPRPPSPASPAQSPAGTPGRASPAQSPASPAQSPAGTPGRASPAQAS 60
 DB 1 MERDSHGASAPRPPSPASPAQSPAGTPGRASPAQSPASPAQSPAGTPGRASPAQAS 60
 QY 61 PACTPPGRASPCR-----ASPAQASPARASPALASLSRSSSSGRSSASASAVTTSPTRYV 115
 DB 61 PACTPPGRASPCRASPAQASPAQASPARASPALASLSRSSSSGRSSASASAVTTSPTRYV 120
 QY 116 LVRAATPVGAIVIRSSPARASAPATATATRESPTSLPKFTWREGQQLPIGCVLLIALV 175
 DB 121 LVRAATPVGAIVIRSSPARASAPATATATRESPTSLPKFTWREGQQLPIGCVLLIALV 180
 QY 176 SLILPFOGHTGIRYKEGRESCKPAVRCGVDCGLKSDDELGCVRPMDKSLTKYS 235
 DB 181 SLILPFOGHTGIRYKEGRESCKPAVRCGVDCGLKSDDELGCVRPMDKSLTKYS 240
 QY 236 GSSHQWMLPICSSNNNDYSEKTCQOLGPFESARTEVAHRDPFANSFILRYNSTIOESLH 295
 DB 241 GSSHQWMLPICSSNNNDYSEKTCQOLGPFESARTEVAHRDPFANSFILRYNSTIOESLH 300
 QY 296 RSECPSPORYSISLOCSHCGLRMTGRIVGALASDSKMPWVSLHFGTTHICGTLIDAQW 355
 DB 301 RSECPSPORYSISLOCSHCGLRMTGRIVGALASDSKMPWVSLHFGTTHICGTLIDAQW 360
 QY 356 VLTAHCFVTVREKVLBSGMKYAGTNSNHLPEASTAIIINSNYDEEDDYIALMRL 415
 DB 361 VLTAHCFVTVREKVLBSGMKYAGTNSNHLPEASTAIIINSNYDEEDDYIALMRL 420
 QY 416 SKPLTLNHIHACLPNMGQTFSLNCTWITGFGKTRTDDTSPFLREVOVNLIDFKKC 475
 DB 421 SKPLTLNHIHACLPNMGQTFSLNCTWITGFGKTRTDDTSPFLREVOVNLIDFKKC 480
 QY 476 NDVLYVDSYLTLPNMGACGLRGRSDSCGDSGGPLVCEQNNRWYLAGVTSWGTGGGQNNK 535
 DB 481 NDVLYVDSYLTLPNMGACGLRGRSDSCGDSGGPLVCEQNNRWYLAGVTSWGTGGGQNNK 540
 QY 536 PGVYTKYTEVLPWTYSKME 554
 DB 541 PGVYTKYTEVLPWTYSKME 559

RESULT 13

ADH17434
 ID ADH17434 standard; protein; 586 AA.
 XX
 AC ADH17434;
 DT 11-MAR-2004 (first entry)
 XX
 DE Human NOV12k protein - SEQ ID 124.
 XX
 KW NOVX; antidiabetic; anorectic; cardiant; hypotensive;
 KW antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
 KW protozoacide; neuroprotective; antiinflammatory;
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
 KW dermatological; antiaesthetic; antipneumonic; metabolic; diabetes;
 KW obesity; infection; anorexia; cancer; cardiovascular; hypertension;
 KW atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's;
 KW epilepsy; immune; osteoarthritis; haemopoietic;
 KW inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
 KW cell differentiation; proliferation; haemopoiesis; wound healing;
 KW angiogenesis; gene therapy; chromosome mapping; tissue typing;
 KW pharmacogenomic; human.
 XX
 OS Homo sapiens.
 XX
 PN MO2003093432-A2.
 XX
 PD 13-NOV-2003.
 XX
 PF 02-MAY-2003; 2003WO-US013690.
 XX
 PR 02-MAY-2002; 2002US-0377321P.
 PR 08-MAY-2002; 2002US-0378730P.
 PR 24-MAY-2002; 2002US-0383075P.
 PR 29-MAY-2002; 2002US-0384044P.
 PR 30-MAY-2002; 2002US-0384215P.
 PR 30-MAY-2002; 2002US-0384296P.
 PR 30-MAY-2002; 2002US-0384327P.
 PR 30-MAY-2002; 2002US-0384352P.
 PR 31-MAY-2002; 2002US-0385211P.
 PR 02-JUL-2002; 2002US-0393333P.
 PR 09-AUG-2002; 2002US-0402154P.
 PR 09-AUG-2002; 2002US-0402171P.
 PR 09-AUG-2002; 2002US-0402204P.
 PR 09-AUG-2002; 2002US-0402205P.
 PR 22-AUG-2002; 2002US-0405175P.
 PR 27-AUG-2002; 2002US-0406129P.
 PR 23-SEP-2002; 2002US-0412954P.
 PR 30-SEP-2002; 2002US-0414975P.
 PR 07-OCT-2002; 2002US-0416651P.
 PR 24-OCT-2002; 2002US-0420851P.
 PR 31-OCT-2002; 2002US-0422547P.
 PR 01-MAY-2003; 2003US-00428275.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alvarez E, Anderson DM, Boldog FL, Catterton E, Edinger SR,
 PI Fernandes ER, Gerlach VL, Gorman L, Grosse WM, Guo X, Ji W,
 PI Kekuda R, Li L, MacDougall JR, Padigaru M, Patnurejan M,
 PI Peterson JD, Rastelli L, Shinkets RA, Spytek KA, Stone DJ,
 PI Vernet CM, Voss EZ, Zhong M;
 XX
 DR WPI: 2004-053040/05.
 DR N-PSDB: ADH17433.
 XX
 PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
 PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
 PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 124; 478bp; English.
 XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates antidiabetic, anorectic,

CC cardiometabolic, hypotensive, antihypertensive, anorectic, vinorelbine,
CC antibacterial, fungicide, protocoarctate, neurotropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteoprotective, antitardic,
CC antiinflammatory, dermatological, antihistaminic and antipneumatic
CC activities. The polypeptides, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.

SQ Sequence 586 AA;

Query Match	98.3%	Score 2948.5;	DB 8;	Length 586;
Best Local Similarity	99.1%	Pred. No. 1.8e-185;		
Matches 554; Conservative	0;	Mismatches 0;	Indels 5;	Gaps 14.

QY	I	1	MEMRSHGNASAPARTPSAASAPQAS	PACTPPERRAS	PAQAS	SPQAS	PACTPPERRAS	PAQAS	60
Db	1		MEMRSHGNASAPARTPSAASAPQAS	PACTPPERRAS	PAQAS	SPQAS	PACTPPERRAS	PAQAS	60
QY	61		PACTPPGRASFOR-----ASPAQAS	PAPAS	PALAS	LSRSS	SGRSS	SSARSASATTSPTRY	115
Db	61		PACTPPGRASFORASPAQASPAQAS	PAPAS	PALAS	LSRSS	SGRSS	SSARSASATTSPTRY	120
QY	116		LVRATPVGAVDIRSSPARSAPATRAS	PSTSLPKFTWREGOKL	PLIGVLLIALV				175
Db	121		LVRATPVGAVDIRSSPARSAPATRAS	PSTSLPKFTWREGOKL	PLIGVLLIALV				180
QY	176		SLILLFOWOGHTGTRYKEORRESC	KHAVRCGVYDCKL	KSDDELCAV	FMDM	SLILKITS		235
Db	181		SLILLFOWOGHTGTRYKEORRESC	KHAVRCGVYDCKL	KSDDELCAV	FMDM	SLILKITS		240
QY	236		GSSHOMLPGICSSNMWDSYSEKTC	COOLGFEASARTTEVAHRDPANS	FSILIRVNSTIQESLH				295
Db	241		GSSHOMLPGICSSNMWDSYSEKTC	COOLGFEASARTTEVAHRDPANS	FSILIRVNSTIQESLH				300
QY	296		RSECPSORYISLQSGHCGLRAMTGR	IIVGALASDSKWPQVSLHFGTTHI	CGGLIDANQ				355
Db	301		RSECPSORYISLQSGHCGLRAMTGR	IIVGALASDSKWPQVSLHFGTTHI	CGGLIDANQ				360
QY	356		VLTAAHCFVYTRERKVLBGMKYVY	AGASNLHOLPEASISLEI	IINSYTBEDDYDIALMRL				415
Db	361		VLTAAHCFVYTRERKVLBGMKYVY	AGASNLHOLPEASISLEI	IINSYTBEDDYDIALMRL				420
QY	416		SKPLTLSAHIAPACLPMHGOTFSL	NETCMTIGFGKTRTENDPTS	SPLREVOYNLIDFPKC				475
Db	421		SKPLTLSAHIAPACLPMHGOTFSL	NETCMTIGFGKTRTENDPTS	SPLREVOYNLIDFPKC				480
QY	476		NDYLVYDSYLLPRMWCAGDLRG	GRDSCQGDSDGPLVCEONNRMY	LAGVTSWGTCGQGRNK				535
Db	481		NDYLVYDSYLLPRMWCAGDLRG	GRDSCQGDSDGPLVCEONNRMY	LAGVTSWGTCGQGRNK				540
QY	536		PGVYTKYTEVLPWITSKME	554					
Db	541		PGVYTKYTEVLPWITSKME	559					

AC	ADH17430;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	Human NOV121 protein - SEQ ID 120.
XX	
XX	NOVY; antidiabetic; anorectic; cardiant; hypotensive;
KM	antiartherosclerotic; anorectic; vituicide; antibacterial; fungicide;
KM	prozoacetic; nootropic; neuroprotective; antiparkinsonian;
KM	anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
KM	demacoulant; antiseptic; antitumoric; metabolic; diabetes;
KM	obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KM	atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's
KM	epilepsy; immune; osteoarthritis; hemopoietic;
KM	inflammatory skin disorder; asthma; dyslipidaemia; neurogenesis;
KM	cell differentiation; proliferation; haemopoiesis; wound healing;
KM	angiogenesis; gene therapy; chromosome mapping; tissue typing;
XX	pharmacogenomic; human.

OS Homo sapiens.

PN WO2003093432-A2

PD 13-NOV-2003.

PF 02-MAY-2003; 2003WO-US013690.

PR	28-MAY-2002	2002US-0377321P
PR	08-MAY-2002	2002US-0378730P
PR	24-MAY-2002	2002US-0383075P
PR	29-MAY-2002	2002US-0384415P
PR	30-MAY-2002	2002US-0384541P
PR	30-MAY-2002	2002US-0384936P
PR	30-MAY-2002	2002US-0384979P
PR	30-MAY-2002	2002US-0384932P
PR	30-MAY-2002	2002US-0384527P
PR	30-MAY-2002	2002US-0384527P
PR	31-MAY-2002	2002US-0385131P
PR	02-JUL-2002	2002US-0393333P
PR	02-JUL-2002	2002US-0402154P
PR	09-AUG-2002	2002US-0402071P
PR	09-AUG-2002	2002US-0402024P
PR	09-AUG-2002	2002US-0402024P
PR	09-AUG-2002	2002US-0402024P
PR	22-AUG-2002	2002US-0405175P
PR	22-AUG-2002	2002US-0405175P
PR	27-AUG-2002	2002US-0406129P
PR	23-SEP-2002	2002US-0412854P
PR	30-SEP-2002	2002US-0414875P
PR	07-OCT-2002	2002US-0418611P
PR	24-OCT-2002	2002US-0420561P
PR	31-OCT-2002	2002US-0422847P
PR	01-MAY-2003	2003US-0428275P

PA (CURA-) CURAGEN CORP.

PI Alvarez E, Anderson DW, Boldog FL, Catterton E, Edinger SE
PI Fernandes ER, Gerlach VL, Gorman L, Grose WM, Guo X, Ji
PI Kekuda R, Li L, Macdonald JR, Padigan M, Paturajan M,
PI Peterson JD, Rastelli L, Shinkens RA, Spytek KA, Stone DJ,
PI Vernet CM, Voss EZ, Zhong M;

DR WPI; 2004-053040/05.

DR N-PSDB; ADH17429.

PT New isolated NOVX polypeptide, useful for preventing, diagnosing or
 PT treating NOVX-associated disorders, e.g. osteoarthritis, obesity,
 PT atherosclerosis, cancer, Parkinson's disease, asthma, or infections
 XX
 PS Claim 1; SEQ ID NO 120; 478bp; English.

The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates antidiabetic, anorectic, cardiatic, hypotensive, antihypertensive, anorectic, vitnucide, antibacterial, fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian, anticonvulsant, osteopathic, antiallergic,

anti-inflammatory, dermatological antihistaminic and antipruritic activities. The polypeptides, nucleic acid molecules and antibiotics may be useful in the manufacture of a medicament for treating metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders such as osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and various types of dyslipidaemia. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit neurogenesis, cell differentiation, cell proliferation, haemopoiesis, wound healing and angiogenesis, in gene therapy and the in generation of antibodies that bind immunospecifically to NOxv substances for use in therapeutic or diagnostic methods. The nucleic acids may be further used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The current sequence is that of the human NOxv protein of the invention.

SQ Sequence 593 AA;

Query Match	98.3%;	Score 2948.5;	DB 8;	Length 593;
Best Local Similarity	99.1%;	Pred. No. 1.8e-185;		
Matches 554;	Conservative 0;	Mismatches 0;	Indels 5;	Gaps 1.

Qy	I	MERSHGNASAPRTSASAPQASAPACPPRRASAPQASAPQASAPACPPRRASAPQAS	60
Db	5	MERSHGNASAPRTSASAPQASAPACPPRRASAPQASAPQASAPACPPRRASAPQAS	64
Qy	61	PACPPRRASAPR-----ASPAQASAPARASPALASISRSSSSRSASARSASVTSPTTRY	115
Db	65	PACPPRRASAPRASAPQASAPQASAPARASPALASISRSSSSRSASARSASVTSPTTRY	124
Qy	116	LVRATPVGAVPIRSSAPASAPATRASPTSGTSLPKFTWREGOKOLPLIGCVLLIALVY	175
Db	125	LVRATPVGAVPIRSSAPASAPATRASPTSGTSLPKFTWREGOKOLPLIGCVLLIALVY	184
Qy	176	SLIILFQWOGHTGIRYKEQRESCKHANCQVVDCKLKSDELGVRRDPMKSLIKITS	235
Db	185	SLIILFQWOGHTGIRYKEQRESCKHANCQVVDCKLKSDELGVRRDPMKSLIKITS	244
Qy	236	GSSHOMLEICSSNMWDSYSEKTCQOLGFEASARTEVAHRDPANSPSILRYNSTIOESLH	295
Db	245	GSSHOMLEICSSNMWDSYSEKTCQOLGFEASARTEVAHRDPANSPSILRYNSTIOESLH	304
Qy	296	RSECBQRYISLQCSHCGLRAMTGRIVGGALASDSKMPQVSLHFQTTIICGTLIDAQW	355
Db	305	RSECBQRYISLQCSHCGLRAMTGRIVGGALASDSKMPQVSLHFQTTIICGTLIDAQW	364
Qy	356	VLTAAHCFVTRKTELEGVKTVAGTSNLHQLPEAASIAEIIINSYNTDEEDYDIAMRL	415
Db	365	VLTAAHCFVTRKTELEGVKTVAGTSNLHQLPEAASIAEIIINSYNTDEEDYDIAMRL	424
Qy	416	SKPLTLSAHIPACPLMHGQFSLNETCMTISFGKRETRDDKTSPLREVOYNLIDFKKC	475
Db	425	SKPLTLSAHIPACPLMHGQFSLNETCMTISFGKRETRDDKTSPLREVOYNLIDFKKC	484
Qy	476	NDYLVYDSYLTFRNMWCAGDLRGGRSQQGDSGAPLVCBQNNRMVYLAGVTSWGTGCGQRNK	535
Db	485	NDYLVYDSYLTFRNMWCAGDLRGGRSQQGDSGAPLVCBQNNRMVYLAGVTSWGTGCGQRNK	544
Qy	536	PGVYTKVTEVLPWYISKME	554
Db	545	PGVYTKVTEVLPWYISKME	563

RESULT 15

ID ADH17452 standard; protein; 586 AA.

AC ADH17452;

DT 11-MAR-2004 (First entry)

XX	Human NOV12t protein - SEQ ID 142
DE	

KM NOVA; antidiabetic; anorectic; virulent; hypocensative;
KM antiarteriosclerotic; anorectic; virulent; antifungal;
KM protozoacide; nootropic; neuroprotective; antiparkinsonian;
KM anticonvulsant; osteopathic; antitachytic; antinflammatory;
KM dermatological; antiasthmatic; antilipemic; metabolic; diabetes;
KM obesity; infection; anorexia; cancer; cardiovascular; hypertension;
KM atherosclerosis; neurodegenerative; Alzheimer's disease; Parkinson's
KM epilepsy; immune; osteoarthritis; hemiplegia;
KM inflammatory skin disorder; asthma; dypliodamia; neurogenesis;
KM cell differentiation; proliferation; hemopoiesis; wound healing;
KM angiogenesis; gene therapy; chromosome mapping; tissue typing;
KM pharmacogenomic; human; SNP; single nucleotide polymorphism;

OS Homo sapiens.

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT Misc-difference 401

single nucleotide polymorphism (SNP) "

PN W02003093432-A2

PD 13-NOV-2003.

PF 02-MAY-2003; 2003WO-US013690.

PR	02-MAY-2002	2002US-0377321P
PR	08-MAY-2002	2002US-0378730P
PR	24-MAY-2002	2002US-0383075P
PR	29-MAY-2002	2002US-0384044P
PR	30-MAY-2002	2002US-0384215P
PR	30-MAY-2002	2002US-0384296P
PR	30-MAY-2002	2002US-0384979P
PR	30-MAY-2002	2002US-0384327P
PR	30-MAY-2002	2002US-0384521P
PR	31-MAY-2002	2002US-0384521P
PR	02-JUN-2002	2002US-0393333P
PR	09-AUG-2002	2002US-0402154P
PR	09-AUG-2002	2002US-0402171P
PR	09-AUG-2002	2002US-0402204P
PR	09-AUG-2002	2002US-0402205P
PR	22-AUG-2002	2002US-0406175P
PR	27-AUG-2002	2002US-0406129P
PR	23-SEP-2002	2002US-0412954P
PR	30-SEP-2002	2002US-0414975P
PR	07-OCT-2002	2002US-0416611P
PR	24-OCT-2002	2002US-04208651P
PR	31-OCT-2002	2002US-0422547P
PR	01-MAY-2003	2003US-00428275

PA (CURA-) CURAGEN CORP.

PI Alvarez E, Anderson DW, Boldog FL, Catterton E, Edinger SR,
PI Fernandes ER, Gerlach VL, Gorman L, Grosse NM, Guo X, Ji W
PI Kekuda R, Li L, MacDougall JR, Padigar M, Paturajan M,
PI Peterson JD, Rastelli L, Shinkets RA, Spytek KA, Stone DJ,
PI Vernet CAM, Voss EZ, Zhong M,

DR WPI; 2004-053040/05.

DR N-PSDB; ADH17451.

PT New isolated NOXV polypeptide, useful for preventing, diagnosing or treating NOXV-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections

PS Claim 1; SEQ ID NO 142; 478bp; English.

CC The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates antidiabetic, anorectic,
CC cardiatic, hypotensive, antarteriosclerotic, anorectic, virucide,

CC antibacterial, fungicide, protozoacide, nootropic, neuroprotective,
CC antiparkinsonian, anticonvulsant, osteopathic, antiarthritis,
CC antiinflammatory, dermatological, antiaesthetic and antilipemic
CC activities. The polypeptide, nucleic acid molecules and antibodies may
CC be useful in the manufacture of a medicament for treating metabolic
CC disorders, diabetes, obesity, infectious diseases (viral, bacterial,
CC fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular
CC diseases including hypertension and atherosclerosis, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune
CC disorders such as osteoarthritis, haemopoietic disorders, inflammatory
CC skin disorders, asthma and various types of dyslipidaemia. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit neurogenesis, cell
CC differentiation, cell proliferation, haemopoiesis, wound healing and
CC angiogenesis, in gene therapy and the in generation of antibodies that
CC bind immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids may be further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine and pharmacogenomics. The current sequence is that of the human
CC NOVX protein of the invention.
CC
XX

SQ Sequence 586 AA;

Query Match 98.1%; Score 2943.5; DB 8; Length 586;
Best Local Similarity 98.9%; Pred. No. 3.9e-185;
Matches 553; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 MERDSHGNSAPARTSPASAPQASPAQSPAGPPGRASPAQASPAQSPAGPPGRASPAQAS 60
DB 1 MERDSHGNSAPARTSPASAPQASPAQSPAGPPGRASPAQASPAQSPAGPPGRASPAQAS 60
QY 61 PAGTPPGRASPGR-----ASPAQASPARASPALASISRSSSGRSSSARSASVTTSPTRY 115
DB 61 PAGTPPGRASPGRASPAQASPAQASPARASPALASISRSSSGRSSSARSASVTTSPTRY 120
QY 116 LVRATPVGAVPFRSSPARASAPATRTARSPTGSLPKFTWRBQKQPLIGCVLILIALVY 175
DB 121 LVRATPVGAVPFRSSPARASAPATRTARSPTGSLPKFTWRBQKQPLIGCVLILIALVY 180
QY 176 SLIIIFQFMOGHITGRYRKORSCPKHAIRCVDVDCILKSDCLGCVFPMQDKSLIKIYS 235
DB 181 SLIIIFQFMOGHITGRYRKORSCPKHAIRCVDVDCILKSDCLGCVFPMQDKSLIKIYS 240
QY 236 GSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFILRYNSTIOESLH 295
DB 241 GSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFILRYNSTIOESLH 300
QY 296 RSECPBQRYISIQCSHCGRAMTGRIVGALASDSKPMQVSLHFGTHICGTLIDAW 355
DB 301 RSECPBQRYISIQCSHCGRAMTGRIVGALASDSKPMQVSLHFGTHICGTLIDAW 360
QY 366 VLTAAHCFVTRBKYLEGKVKVAGTSNLHQLPEASIAIITNSNYTDEEDDYIALML 415
DB 361 VLTAAHCFVTRBKYLEGKVKVAGTSNLHQLPEASIAIITNSNYTDEEDDYIALML 420
QY 416 SKPLTLSAHIHACLPMHGQTFSLNETCMITGFKTRSTDXTSPFLREVQVNLIDFKKC 475
DB 421 SKPLTLSAHIHACLPMHGQTFSLNETCMITGFKTRSTDXTSPFLREVQVNLIDFKKC 480
QY 476 NDYLYVDSYLTFRMMCAQDLRGRSDSCQDSGGLVCEONNRWYLAGTWSGTGCGQRNK 535
DB 481 NDYLYVDSYLTFRMMCAQDLRGRSDSCQDSGGLVCEONNRWYLAGTWSGTGCGQRNK 540
QY 536 PGVYTKVTEVLPMIYSKME 554
DB 541 PGVYTKVTEVLPMIYSKME 559

Search completed: September 16, 2006, 03:08:13
Job time : 201 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:08:31 ; Search time 43 Seconds

(without alignments)
1257.531 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASAPRTPSAGAS.....TEVLPWYKMESEVPRKRS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	611.5	20.4	417	1 S00845	hepsin (EC 3.4.21.
2	592.5	19.8	416	1 S33777	hepsin (EC 3.4.21.
3	575	19.2	1019	1 A56318	enteropeptidase (E
4	574	19.1	638	1 KQMSPL	plasma kallikrein
5	572.5	19.1	638	1 KQMSPL	plasma kallikrein
6	568.5	19.0	855	2 J07731	membrane-bound arg
7	565.5	18.9	638	1 KQMSPL	plasma kallikrein
8	554	18.5	1034	1 A53663	enteropeptidase (E
9	549	18.3	1113	2 J03315	low-density lipopr
10	548.5	18.3	1035	1 A43090	enteropeptidase (E
11	536	17.9	761	2 J05759	brain-specific ser
12	532	17.7	625	1 KQMSPL	coagulation factor
13	514	17.1	812	1 PLEBO	plasma (EC 3.4.21
14	510.5	17.0	343	1 A57014	proctasin (EC 3.4.
15	508	16.9	274	2 J04171	trypsin (EC 3.4.2
16	506.5	16.9	455	2 A61545	plasma (EC 3.4.21
17	505	16.8	276	2 A38654	maest cell) proteina
18	495	16.5	460	2 B61545	plasma (EC 3.4.21
19	495	16.5	460	2 B61545	plasma (EC 3.4.21
20	493.5	16.5	810	1 PLEBO	plasma (EC 3.4.21
21	491	16.4	270	2 S56160	maest cell) trypsin
22	488.5	16.3	273	2 A47246	trypsin (EC 3.4.2
23	480	16.0	785	2 C35863	serine proteinase
24	480	16.0	785	2 A47547	plasma (EC 3.4.21
25	480	16.0	790	1 PLEBO	plasma (EC 3.4.21
26	479	16.0	275	2 A35863	trypsin (EC 3.4.2
27	479	16.0	275	2 A35863	trypsin (EC 3.4.2
28	478	15.9	275	2 B35863	trypsin (EC 3.4.2
29	477.5	15.9	418	2 A37344	acrosin (EC 3.4.21

30	477	15.9	275	2 A32410	trypsin (EC 3.4.2
31	473	15.8	437	2 S18407	acrosin (EC 3.4.21
32	472.5	15.8	810	2 B30848	plasma (EC 3.4.21
33	471.5	15.7	4548	1 S00657	apoptotain(a) (EC
34	470	15.7	436	2 UK0172	acrosin (EC 3.4.21
35	468.5	15.6	1524	2 J30337	polyprotein - Alti
36	466.5	15.6	1420	2 A32869	apolipoprotein (a)
37	466	15.5	421	2 S29599	acrosin (EC 3.4.21
38	466	15.5	655	1 A46688	hepatocyte growth
39	457.5	15.3	274	2 A45754	trypsin (EC 3.4.2
40	457	15.2	431	2 S47538	acrosin (EC 3.4.21
41	456	15.2	810	2 I46260	plasma (EC 3.4.21
42	453.5	15.1	421	1 S11674	acrosin (EC 3.4.21
43	447	14.9	434	1 A35005	u-plasminogen acti
44	438	14.6	237	2 S68702	trypsin (EC 3.4.2
45	435	14.5	562	1 UKHUT	t-plasminogen acti

ALIGNMENTS

RESULT 1

S00845
hepsin (EC 3.4.21.-) - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004
C:Accession: S00845
R:Leysen, S.P.; Leeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
Biochemistry 27, 1067-1074, 1988
A:Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane dome
A:Reference number: S00845; MUID:88209431; PMID:2835076
A:Accession: S00845
A:Molecule type: mRNA
A:Residues: 1-417 <LE>
A:Cross-References: UNIPROT:P05981; UNIPARC:UPI000003FE67; EMBL:X07732; NID:G32063; PIDN:
C:Genetics:
A:Gene: GDB:HPN; TMPSR1; hepsin
A:Cross-References: GDB:135685; OMIM:142440
A:Map position: 19q11-19q13.2
A:Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F:23-45/Domain: transmembrane #status predicted <TM>
F:163-400/Domain: trypsin homology <TRY>
F:188-204,291-359,322-338,349-381/Denulfide bonds: #status predicted
F:203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 20.4%; Score 611.5; DB 1; Length 417;
Best Local Similarity 36.0%; Pred. No. 2.3e-26;
Matches 134; Conservative 62; Mismatches 147; Indels 29; Gaps 9;

QY	209	VVDCKLXGDE--LGCVRPDMKSLIKYSGSSHOWLPICSSNMWDSYSEKTCQQLGFESA	266
DB	39	IVAVLTKSDQEPRLVQVSSADARLMDVDTKGTGTRLLCSSNSNRNVAGLSCGEWGLRA	98
QY	267	HRTTEVARD-----FANSFSLRYNSTIOESLHRSBCPSQRYSLQCSHGLRAM	317
DB	99	LTHSELVDRTGANGSTGFCVDEBGRLPHTQRLLEVISVCCPCRRFLAICODCGRRRL	158
QY	318	TGRIVGALASDSKMPQVSLHFGTHIICGTLIDAQWVTLAARCFVTRKYLEGKV	376
DB	159	PVDRIYVGRDRLSLGMPQVSLRYDGAHLCGSLSLGDMVTLAHC-PPERNRVLSRRRV	217
QY	377	YAGT---SNLHQLPEAASIAEIIINSNY-----TDEDDVDIALMRSLKPTLSAHHP	427
DB	218	PAGAVQAQSPHGL-QLGVAQVYHGGILPRDPNSENNDIALVHSSPLPLEYIOP	275
QY	428	ACLPMHGGTFSINMETCWTGFGKTRTDKTSPLREVOVNLIDFKKNDVLYVDSYLTLP	487
DB	276	VCLPAAGQALVDGKICTVGTGNGNTQYGGQAG-VLQEARVPLISNDVCGADFYNQIKP	334
QY	488	RMKMGDRLGRDSCQSGSGPVE-----QNNRYTLAGVTSWGCQGRNKPQVYTVT	543
DB	335	KMFAGYPEGGIDACQSGSGPVECDSDISRTPRRLCGIVSWGICALAQPGVYTVTS	394

Oy 544 EULPMYKMS 555
 Db 395 DFERWIFQAIKT 406

RESULT 2

S33777
 hepsin (BC 3.4.21.-) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_rev100 06-Jan-1995 #text_change 05-Oct-2004
 C:Accession: S33777; S32013
 R:Farley, D.; Raymond, F.; Nick, H.
 Biochim. Biophys. Acta 1173, 350-352, 1993
 A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
 A:Reference number: S33777; MUID:93305733; PMID:8318546
 A:Accession: S33777

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <FAR>
 A:Cross-references: UNIPROT:005511; UNIPARC:UPI000048CD9; EMBL:X70900; NID:g57928; PIDN:
 F:12-44/Domain: hydrolase; liver; serine proteinase; transmembrane protein
 F:162-389/Domain: trypsin homology <TRY>
 F:187-203,280-358,321-337,348-380/Dissulfide bonds: #status predicted
 F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 19.8%; Score 592.5; DB 1; Length 416;
 Best Local Similarity 36.3%; Pred. No. 2,56-25;
 Matches 138; Conservative 59; Mismatches 138; Indels 45; Gaps 12;

Oy 209 VVDCXKLSDE--LGCVRPMDKSLIKIYSGSHQWLPICSSNMNDYSEKTCQQLGFPSA 266
 Db 38 IVTLLRSDQDFLYQVQLSPGSRLLVLDKTEGTRLLCSSRSNARVAGLCCEWGFARA 97
 Oy 267 HRTTEVAHRDF-----NSFS-----ILKNSTIOESLHRSCEQSQRYISLQCSHC 312
 Db 98 -----LHASELDVRTAGNAGTSGFPCVDEGGPLAQLLDVLSVDCRGGRPLTATCCDC 152
 Oy 313 GLRAM-TGRIVGALASDKMPQVSLHFGTTHICGTLIDQWVLTAAHCFVTRKYL 371
 Db 153 GRRLPVRIVYGGQSSSLGRMPQVSLRYDGTLLCGSLLSGDWVLTAAHCFVTRKYL 211
 Oy 372 EGMKVYAG---TSNHLQPEAA-SIAETIINSY-----TDEEDYDIALMLSLRPL 419
 Db 212 SRWRFAAGAVARTS-----PAAVQLGVAIVYHGGLPFRDPTIDENSNDIALVHLSSSL 266
 Oy 420 TLSAHHHACLPNGQGTSLNBTGWTGFGKTRERDDTSPLRQVQVNLIDFKKQNDYL 479
 Db 267 PLTEIIOVCLPAAQOALVDGKVCVTGNGMT-QFYGOAAVVLQARVPILISNEVCNSPD 325
 Oy 480 VYDSYLTFRMWCAGDLRGGRDSCQDSSGAPLVCEQ---NNRWYLAGVTSNGTCCGQRNK 535
 Db 326 FYGNQIKRKMFCAGPRBGGIDACQDSSGHHFVCEBRIGTSRMRLCGIVSNGTGCALARK 385
 Oy 536 PGVYTKVTEVLPWYKMS 555
 Db 386 PGVYTKVIDFERWIFQAIKT 405

RESULT 3

A56318
 enteropeptidase (BC 3.4.21.9) precursor [validated] - human
 N:Alternate names: enterokinase
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1995 #sequence_rev100 09-Aug-1996 #text_change 09-Jul-2004
 C:Accession: A56318; B43090
 R:Kitamoto, Y.; Velle, R.A.; Donis-Keller, H.; Sadler, J.E.
 Biochemistry 34, 4562-4568, 1995
 A:Title: cDNA sequence and chromosomal localization of human enterokinase, the proteolytic
 A:Reference number: A56318; MUID:95234679; PMID:7718557
 A:Accession: A56318
 A:Molecule type: mRNA
 A:Residues: 1-1019 <KIT>

A:Cross-references: UNIPROT:P98073; UNIPARC:UPI000003FE65; GB:U098660; NID:g746412; PIDN:
 R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.
 Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
 A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed
 A:Reference number: A43090; MUID:94329561; PMID:8052624
 A:Accession: B43090

A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 749-1019 <K12>
 A:Cross-references: UNIPARC:UPI0000172B0B; GB:U098660
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 (located below) or with amino-terminal myristoylation of the heavy chain.
 A:Gene: GDB:PR857
 A:Cross-references: GDB:384083; OMIM:226200
 A:Map position: 21q21-21q21
 C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
 ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involve
 ducts.

C:Function:
 A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolase cascade
 C:Superfamily: enteropeptidase; Ctr/Cts repeat homology; LDL receptor ligand-binding rep
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
 F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
 F:12-38/Domain: transmembrane #status predicted <TMN>
 F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:342-504/Domain: MAM homology <MAM>
 F:526-631/Domain: Ctr/Cts repeat homology <CTR>
 F:643-677/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:768-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
 F:785-1019/Product: enteropeptidase light chain #status predicted <LCH>
 F:116-147,179,328,335,388,440,470,503,534,630,682,706,725,848,887,909,949/Binding site: C
 F:772-886,910-926,910-977,941-956,967-995/Dissulfide bonds: #status predicted
 F:823,876,971/Active site: His, Asp, Ser #status predicted

Query Match 19.2%; Score 575; DB 1; Length 1019;
 Best Local Similarity 33.3%; Pred. No. 5,56-24;
 Matches 133; Conservative 58; Mismatches 156; Indels 52; Gaps 12;

Oy 187 HTGI-----RYKQRESCKRAVCDGVVDCXKLSDELQCFRPMDSKLIKYSQSSH 239
 Db 636 HLGIPKPKADHPCKKGECPVLNLCQGLHCDGSDADCVRF-----FNGTTN 686
 Oy 240 Q-----WPCSSNMNDYSEKTCQQLGFPSAHRTEVAHRDFANSFILRYNS 288
 Db 687 NGLVRRIGSIWTTACAKEMTTQISNDVCGLLGSGSNSKPIFSTD---GGEFVKLT 743
 Oy 289 TIOESL---HRSECPQRYISLQCSH---CGLRA---MTGRIVGALASDKMPQVSLH 339
 Db 744 APDGLHILTPSQQLQSLRLQCNHKSCKKLAADITPPIVGSNAKESGAMPVVLVLY 803
 Oy 340 FGTTHICGTLIDQWVLTAAHCFVTRKYLEGMKVYAGTSLNHLQPEAAS-----I 392
 Db 804 YGGRLLCGASLVSSDWSAHCYV-GRNLEPSKMTILG---LHMKNLTPOTVPLRI 859
 Oy 393 AEITINSYVTEEDDYDIALMLSLRPLTSAHHHACLPNGQGTSLNBTGWTGFGKTR 452
 Db 860 DEIVINPHYARRKNDNDIAMHLEFKVNTYDIPICLPEENQVFPFRKNSISAGW- TV 918
 Oy 453 ETDKTSPEFLAEVQVNLIDFKKQNDYLVDSYLTFRMWCAGDLRGGRDSCQDSSGAPLV 512
 Db 919 VYGGTANILQDAVPLSNRCCQQNP-EVNITENNICAGYEGGIDSCQDSSGAPLMC 977
 Oy 513 EKNRWYLAGVTSNGTCCGQRNKRGVYTKVTEVLPWYKMS 551
 Db 978 QENRWYLAGVTSNGTCCGQRNKRGVYTKVTEVLPWYKMS 551

RESULT 4

K0MSPL
 kallikrein (BC 3.4.21.34) precursor - mouse

C:Species: Mus musculus (house mouse)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004
 C:Accession: A36557
 R:Setdian, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaudien, G.; Brachpapa, L.; Rochmon
 DNA Cell Biol. 9, 737-746, 1990
 A:Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and compariso
 A:Reference number: A36557; MUID:91090844; PMID:2264928
 A:Accession: A36557
 A:Molecule type: mRNA
 A:Residues: 1-638 <SEI>
 A:Cross-references: UNIPROT:P26262; UNIPARC:UPI00000277BC; GB:M56588; NID:9200358; PIDN:
 A:Note: part of this sequence, including the amino ends of both the heavy and light cha
 C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
 C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li
 C:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-390/Product: plasma kallikrein heavy chain #status experimental <HCH>
 F:20-109/Domains: apple repeat <AP1>
 F:110-139/Domains: apple repeat <AP2>
 F:200-289/Domains: apple repeat <AP3>
 F:291-380/Domains: apple repeat <AP4>
 F:391-638/Product: plasma kallikrein light chain #status experimental <LCH>
 F:391-621/Domains: trypsin homology <TRY>
 F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322
 F:127,215,308,396,494/Binding site: carbonyl (Asn) (covalent) #status predicted
 F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 574; DB 1; Length 638;
 Best Local Similarity 33.7%; Pred. No. 3,9e-24;
 Matches 131; Conservative 61; Mismatches 95; Indels 102; Gaps 15;

```

OY 196 RESCPKAVRCDDGVV-----DCKLKSDLEG--VRFDMKSLIKIYSSGHQWLPIC 246
DB 319 QETCTK-TIRCOFPTYSLLPDCK---EEGCKSLRLSTDGSPRTIYTG----- 363
OY 247 SNNMDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLHSECPGQRIS 306
DB 364 MQGSSGYSRLRLCKLV--DSPDCTYKI----- 387
OY 307 LQSGHGLRAMTGRVVGALASDSKMPQVSLH---FGTHICGTLIDAQWVLTAAICF 363
DB 388 -----NARIVGNTASLSEMPQVSLQVLTQVTHLCGSSIIIGRWVLTAAICF 436
OY 364 FVIREKYLEG-----WKVYAGTSLHQLPE---AASIAEIIINSNYTDEEDVDIALM 414
DB 437 -----DGLPYDPVWRIYSGIISLSEITKETPESRIKEIIHQEYKVSSENYDIALIK 488
OY 415 LSKPLTSAHIHPACLPNMGQTSINETCWITGPKTRTDDKTSPLREVOVNLIDPK 474
DB 489 LOTPLNTEPQKPICLPGRADNTNITYNQWYGMWYIYE--QGETNIILOKATIPLPVPEE 547
OY 475 C-----NDVLVDSYLTTPMMACGLRGRDSCQSGGPGPLVCEONRMYLVAGVTSWGTCG 530
DB 548 COKKTRDYVIT-----NKKQWICAGYKEGGTDACKGDSGGLPKKHSGRQVLGITSWGECG 602
OY 531 GQRNKPQVYTKYTEVLPWYISKME--SEVR 558
DB 603 GRKQPGVYTKVSEYMDWILEKTQSSDVR 631

```

RESULT 5
 KOHUP
 Plasma kallikrein (EC 3.4.21.34) precursor - human
 N:Alternate names: kininogenin; plasma prekallikrein
 C:Species: Homo sapiens (man)
 C>Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Oct-2004
 C:Accession: A00921; A37939
 R:Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
 Biochemistry 25, 2410-2417, 1986
 A:Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four p
 A:Reference number: A00921; MUID:86243359; PMID:3521732

A:Accession: A00921
 A:Molecule type: mRNA
 A:Residues: 1-638 <CHU>
 A:Cross-references: UNIPROT:P03952; UNIPARC:UPI00000008AE; GB:M3143; NID:9190262; PIDN:
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2050-2056, 1991
 A:Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of 1
 A:Reference number: A37939; MUID:91152016; PMID:1998666
 A:Accession: A37939
 A:Molecule type: Protein
 A:Residues: 20-27,40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80,103-113,131-140,141-1
 260-283, 'X', 285,287-291, 'X', 293-295,314-317, 'X', 319-320,321-324, 'X', 329-333,334-339, 'X',
 525,538-551,562, 'X', 564-567,573, 'X', 575-576,578-583, 'X', 585,592-604 <MCA>
 A:Cross-references: UNIPARC:UPI0000172B42; UNIPARC:UPI0000172B43; UNIPARC:UPI0000
 10000172B52; UNIPARC:UPI0000172B53; UNIPARC:UPI0000172B54; UNIPARC:UPI0000172B55; UNIPARC
 10000172B52; UNIPARC:UPI0000172B53; UNIPARC:UPI0000172B54; UNIPARC:UPI0000172B55; UNIPARC
 C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
 C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li
 C:Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal re
 inogen and may also play a role in the renin-angiotensin system by converting prorenin ir
 C:Genetics:
 A:Gene: GDB:KLM3
 A:Cross-references: GDB:127575; OMIM:229000
 A:Map position: 4q35-4q35
 C:Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; inflan
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-638/Product: plasma kallikrein #status predicted <MAT>
 F:20-390/Domains: plasma kallikrein heavy chain #status predicted <HCH>
 F:20-109/Domains: apple repeat <AP1>
 F:110-139/Domains: apple repeat <AP2>
 F:200-289/Domains: apple repeat <AP3>
 F:291-380/Domains: apple repeat <AP4>
 F:391-638/Domains: plasma kallikrein light chain #status predicted <LCH>
 F:391-621/Domains: trypsin homology <TRY>
 F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383-
 F:127,308,396,453,494/Binding site: carbonyl (Asn) (covalent) #status experimental
 F:318-347,340-345/Disulfide bonds: #status predicted
 F:390-391/Cleavage site: Arg-11e (coagulation factor XIIa) #status predicted
 F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 572.5; DB 1; Length 638;
 Best Local Similarity 34.5%; Pred. No. 4.7e-24;
 Matches 132; Conservative 53; Mismatches 101; Indels 97; Gaps 13;

```

OY 196 RESCPKAVRCDDGVV-----DCKLKSDLEG--VRFDMKSLIKI--YSSGHQWLPIC 245
DB 319 QETCTK-MIRCOFPTYSLLPEDC--KEEKCKFLRLSDGSPTRIAYGQSSGYSRLIC 375
OY 246 SNNMDSYSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLHSECPGQRIS 305
DB 376 NTGDNISVCTTYT----- 387
OY 306 SLQSGHGLRAMTGRVVGALASDSKMPQVSLHFGTT---HICGTLIDAQWVLTAAIC 362
DB 388 -----STRIVGNTASLSEMPQVSLQVLTQVTHLCGSSIIIGRWVLTAAIC 435
OY 363 FVIREKYLEG-----WKVYAGTSLHQLPE---AASIAEIIINSNYTDEEDVDIALM 413
DB 436 F-----DGLPYDPVWRIYSGIISLSDIRKDPFQSIKEIIHQYKVSSENYDIALIK 487
OY 414 LSKPLTSAHIHPACLPNMGQTSINETCWITGPKTRTDDKTSPLREVOVNLIDPK 473
DB 488 KLOAPLNTTEPQKPICLPESKSDTSTIYNQWYGMWYIYE--QGETNIILOKATIPLPVPEE 546
OY 474 KNDYLVDSY-LTPMMACGLRGRDSCQSGGPGPLVCEONRMYLVAGVTSWGTCGCGO 532
DB 547 ECQKR--YQDKITQIRWCAHYKGGKADCKGDSGGLVCHNGMRLVGLITSWGECAR 604
OY 533 RNKQGVYTKYTEVLPWYISKMS 555
DB 605 RQPGVYTKVSEYMDWILEKTQS 627

```

```

RESULT 6
UC7731
C:Species: Rattus norvegicus (Norway rat)
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: UC7731; J07775
R/Kishi, K.; Yamazaki, K.; Yasuda, I.; Yahagi, N.; Ichinose, M.; Tsuchiya, Y.; Athauda,
J.; Biochem. 130, 425-430, 2001
A>Title: Characterization of a membrane-bound arginine-specific serine protease from rat
A/Reference number: UC7731; MUID:21421307; PMID:11530019
A/Accession: UC7731
A/Molecule type: mRNA
A/Residues: 1-855 <KIS>
A/Cross-references: UNIPROT:Q9JUI7, UNIPARC:UPI0000088AC0; DBJ:AB049189
A/Experimental source: strain Male, 7-week-old
R/Satom, S.; Yamasaki, Y.; Tsuzuki, S.; Hitomi, Y.; Iwanga, T.; Fushiki, T.
Biochem. Biophys. Res. Commun. 287, 995-1002, 2001
A>Title: A role for membrane-type serine protease (MT-SP1) in intestinal epithelial turn
A/Reference number: J07775; PMID:11573963
A/Contents: Small intestine
A/Accession: J07775
A/Molecule type: mRNA
A/Residues: 1-855 <SAT>
A/Cross-references: UNIPARC:UPI0000088AC0; DBJ:AB037898
A/Comment: This enzyme, an epithelial-derived, type II integral membrane serine protease
of specific proteins or peptides on the brushborder membranes. It also participates in
lial migration and/or cell loss.
C/Genetics:
A/Gene: mt-ep1
A/Map position: basolateral cell surface
A/Superfamily: membrane-bound arginine-specific serine protease
C/Keywords: protein digestion

Query Match 19.0%; Score 568.5; DB 2; Length 855;
Best local similarity 34.5%; Pred. No. 1e-23; Mismatches 133; Indels 81; Gaps 17;
Matches 141; Conservative 54;

199 CPKHAIVRCGVVDCDKLSDELQCVFDMKSLIKYSSSHQWL-----P---ICGS- 247
Db 464 CIRKDLRCDDGMADCPDYSDERHC-----RCNATQIFKCKNFCCKPLFVVCDSV 511

248 -NMNDSYSEKTC-----QQLGFSAHRTTEVAHDPANSPSILRYN 287
Db 512 NDCGGSDBEGSCSPAGSPKCSNGKCLPQSQCKNGKDCGGSDBASCDNNVAVSCTKYT 571

288 STIGESL-----HRSFCPSORXISLQCS-----HCLRAMT--GRVYGALASDSKMP 333
Db 572 YRCQGLCLLNKGNPCCDKK---DCSOGSDEKNCDCGLRSFTKQARVVGATNADGEMP 627

334 MGVSLH-FGTHICGGLIDAQWVLTAAHCF-----FVTRKVLKGMKVYAGTSLHQLP 387
Db 628 MGVSLHAGGGLHLCASLISPDWLSVAHCFQDETFIKSDHTM--WAFPLGL--LDQSK 683

388 EAS-----IAETIINSYTDDEDDYIALMLRSKPLTSLAHAIHACLPHMGQTFSLNE 441
Db 684 RSAGVGOEHKRIITHTSPSFDFTFDYDIALLELEKPARYSTVAVPICLPNTHTVFPAGK 743

442 TCMITGPKRTETDCKTSPFLREVOVNLIDFKCKNDYLVVSYLTPRMWCGDLRGSDS 501
Db 744 AIWYTGWHTYE-GGTGALLIQLKGRIRVINTQTEILL--PQQLTPRMWCVGFLSGVDS 800

502 CGGDSGGL-VCEQNNRWYLAGVTSMTGCGQNRNPGVYTKTEVLPWI 549
Db 801 CGGDSGGLSVEKGRIFQAGVVSMBGCAQRNRKGVYTTIPIVRDVI 849

RESULT 7
KORTPL
C:Species: Rattus norvegicus (Norway rat)
C:Species: Rattus norvegicus (Norway rat)

```

```

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Oct-2004
C/Accession: A39180; A33320; S06851; I53041; S06852
R/Beaubien, G.; Rosinski-Chupin, I.; Mattei, M.G.; Mbikay, M.; Chretien, M.; Seidah, N.G.
Biochemistry 30, 1628-1635, 1991
A>Title: Gene structure and chromosomal localization of plasma kallikrein.
A/Reference number: A39180; MUID:91129236; PMID:1993180
A/Accession: A39180
A/Molecule type: DNA
A/Residues: 1-638 <BEA>
A/Cross-references: UNIPROT:P14272; UNIPARC:UPI000048CC9; GB:J05315
A/Note: The authors translated the codon GAG for residue 81 as Gln
R/Seidah, N.G.; Lidenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
DNA 8, 563-574, 1989
A>Title: The cDNA structure of rat plasma kallikrein.
A/Reference number: A33320; MUID:90091743; PMID:2598771
A/Accession: A33320
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-638 <SEI>
A/Cross-references: UNIPARC:UPI000048CC9; GB:M30282; NID:G205010; PID:AAA41463.1; PID:
A/Note: part of this sequence, including the amino ends of both the heavy and light chain
R/Paglin, J.; Benjannet, S.; Sawyer, N.; Lazure, C.; Chretien, M.; Seidah, N.G.
Biochim. Biophys. Acta 999, 103-110, 1989
A>Title: Rat plasma kallikrein: purification, NH(2)-terminal sequencing and development c
A/Reference number: S06851; MUID:90089457; PMID:2597701
A/Accession: S06851
A/Molecule type: protein
A/Residues: 20-45/391-413 <PAQ>
A/Cross-references: UNIPARC:UPI0000172B57; UNIPARC:UPI0000172B58
R/Seidah, N.G.; Lidenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazure
DNA Cell Biol. 8, 563-574, 1989
A>Title: The cDNA structure of rat plasma kallikrein.
A/Reference number: I53041
A/Accession: I53041
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-638 <RGS>
A/Cross-references: UNIPARC:UPI0000048CC9; GB:M58590; NID:G206721; PID:AAA42069.1; PID:
A/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 145
are linked by one or more disulfide bonds.
C/Genetics:
A/Gene: PK
C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
F1-19/Domin: signal sequence #status predicted <IG>
F120-390/Product: plasma kallikrein heavy chain #status experimental <MAT1>
F120-109/Domin: apple repeat <AP1>
F110-159/Domin: apple repeat <AP2>
F120-289/Domin: apple repeat <AP3>
F1291-380/Domin: apple repeat <AP4>
F1391-621/Product: plasma kallikrein light chain #status experimental <MAT2>
F1391-621/Domin: trypsin homology <TRY>
F121-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,297-375,318-347,322-
F121,215,308,453,459,494/Binding site: carbohydrate (asn) (covalent) #status predicted
F1392/Binding site: carbohydrate (asn) (covalent) #status experimental
F1433,483,578/Active site: His, Asp, Ser #status predicted

Query Match 18.9%; Score 565.5; DB 1; Length 638;
Best local similarity 33.5%; Pred. No. 1.1e-23; Mismatches 95; Indels 107; Gaps 15;
Matches 130; Conservative 56;

196 RESCPKHAIVRCGVV-----DCKLSDELQCVFDMKSLIKI---YSGSSHQWL 243
Db 319 QETCTK-TIRQFTYSLLPDCKAE-----GCKSLSLSTGSPTRITYBAQSS----- 368

244 ICSSNMNDSYSEKTCQQLGFSAHRTTEVAHDPANSPSILRYNSTIOESLHRSCEFSQR 303
Db 369 -----GYSLRLCKV--ESSDCTTKI----- 387

304 YISLQSHGRLAMTGRIVGALASDSKMPQVSLH---FGTHICGGLIDAQWVLTAA 360
Db 388 -----NARIVGSTNSLSGMPQVSLQVLYSQNHMGSGSIIGROMILTAA 433

```

QY 361 HCFVTRKVLG-----WKVYAGTSLHQLPEA---ASIAEIIINSNTDEEDDYDIA 411
 DB 434 HCF-----DGLPYVWMAIYGGILMLSEITNTKTPSSIKELIHOKTKMSSESYDIA 485
 QY 412 LMRSLKPLTSLAHHPACLPNHGQFTSLNETCWTGTFKTRTDKTSPPFLREVQVNLID 471
 DB 486 LIKQPLTANTEPQKPLCLPSKADNTNTLYTNQWLTGWGTYKERGE-TQNILOKATIPLPV 544
 QY 472 FKCC---NDLYVDSYTLTPRMKAGDLRGGRDSCQGSGLPVCQNNRYLAGVTSWG 527
 DB 545 NEBCCKKRYDYVI-----TKQWICAGYKEGGIDACKGDSGPLVCKHSGRWQVGLTSMG 599
 QY 528 TGGCGRNKPGYVTKVTEVLPMTYSQMS 555
 DB 600 EGCAKEQPGVYTKVAYETIMLEKIQS 627

RESULT 8

A53663
 enteropeptidase (BC 3.4.21.9) precursor [validated] - pig
 N/Alternate names: enterokinase
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 07-Oct-1994 #sequence_revision 09-Aug-1996 #text_change 28-Apr-2003
 C/Accession: A53663
 R/Matsuhashima, M.; Ichinose, M.; Yahagi, N.; Kakei, N.; Tsukada, S.; Miki, K.; Kurokawa, U. Biol. Chem. 269, 1976-1998, 1994
 A/Title: Structural characterization of porcine enteropeptidase.
 A/Reference number: A53663; MUID:94327548; PMID:8051081
 A/Accession: A53663
 A/Molecule type: mRNA
 A/Residues: 1-1034 <MNT>
 A/Cross-references: UNIPARC:UPI000017280C; GB:D30799; NID:9505122; PIDN:BA06459.1; PID:
 A/Note: parts of this sequence, including the amino ends of three chains isolated from B
 C/Comment: The mechanism of association with the membrane of the intestinal brush border
 C/Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
 ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involv
 C/Function:
 A/Description: cleaves activation peptide from trypsinogen to produce active trypsin
 A/Pathway: intestinal digestive hydrolase cascade
 C/Superfamily: enteropeptidase; Clr/Cls repeat homology; LDL receptor ligand-binding rep
 C/Keywords: glycoprotein, hydrolase, serine proteinase, transmembrane protein, zymogen
 F/122-38/Domain: transmembrane #status predicted <TM>
 F/152-117/Product: enteropeptidase mini chain #status predicted <MCH>
 F/118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
 F/199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F/357-519/Domain: MAM homology <MAM>
 F/541-646/Domain: Clr/Cls repeat homology <CLR>
 F/658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F/693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
 F/800-1034/Product: enteropeptidase light chain #status predicted <LCH>
 F/800-1029/Domain: trypsin homology <TRY>
 F/116,147,170,194,283,343,350,403,455,518,549,645,697,701,721,740,761,804,863,902,96
 F/787-911,825-841,925-992,956-971,992-1010/Disulfide bonds: #status predicted
 F/840,891,986/Active site: His, Asp, Ser #status predicted

Query Match 18.5%; Score 554; DB 1; Length 1034;
 Best Local Similarity 34.1%; Pred. No. 7.8e-23;

Matches 136; Conservative 56; Mismatches 155; Indels 52; Gaps 14;

QY 187 HTGI-----RYKQRESCPHRAVRCGVVDCUKLSDELGCVRPDMDKSLIKIYSGSS 239
 DB 651 HLGIDPECKEDNFOCEGECVLLVNLCDGFHCKXGSDPAHCVRF-----LNGTAN 701
 QY 240 Q-----WLPICSSNNMDSYSEKTCQOLGFSARHTTEVVAHRDFANSFS--ILRY 286
 DB 702 NSGLVQFRIQSIWHTACANMTTQTSDDVCQLGJGTG-----SMPFFSGGGEFVYL 756
 QY 287 NSTIOESL---HRSECPQRYISLQCSH---CG-----LRAMTRIVGALASDSKMPQWS 337
 DB 757 NTAHPGSLILTASBQCFEDBSLILLCNMHSGCKKQYAGVSPKRTVGANDSRGAMPVVA 816

QY 338 LHFQTHICGGLTIDAQWVLTAAHCFVTRKVLGKRYAG---TSNLHQLPEAAS--I 392
 DB 817 LYVNGQLCGASIVSRDMLVSAACHVY--GRNLEPSKKKALIGLHMTSLNLT--PQIVTRLI 874
 QY 393 AEIINSNTDEEDDYDIALMRSLKPLTSLAHHPACLPNHGQFTSLNETCWTGTFKTRTDKTSPPFLREVQVNLID 471
 DB 875 DEIVNHYNRKRDSDIAMNHLEFKNYTYDYQPLCLPEENQVFPGRICSLAGWGV- 933
 QY 453 ETDKTSPPFLREVQVNLIDPFKCNLYVDSYTLPRMKAGDLRGGRDSCQGSGLPVC 512
 DB 934 IYQSPADILDEADVPLSLSEKCOQNP--EYNTENMKAGYEEGGIDSCGDSGGLPMC 992
 QY 513 EONNRWYLAGVTSWGTGCGRNKPGYVTKVTEVLPMTYS 551
 DB 993 LENNRMLAGVTSFGYCALPNRGVYARVPKFTWIOS 1031

RESULT 9

JE0315
 low-density lipoprotein receptor-related protein - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C/Accession: JE0315
 R/Tomita, Y.; Kim, D.; Magoori, K.; Fujino, T.; Yamamoto, T.T.
 J. Biochem. 124, 784-789, 1998
 A/Title: A novel low-density lipoprotein receptor-related protein with type II membrane
 A/Reference number: JE0315; MUID:9842596; PMID:9756624
 A/Accession: JE0315
 A/Status: Preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1113 <TOM>
 A/Cross-references: UNIPROT:Q92319; UNIPARC:UPI000029874; DBPJ:AB013874; NID:93869144; I
 C/Superfamily: mouse low-density lipoprotein receptor-related protein; LDL receptor ligand
 F/337-372/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F/374-408/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F/410-445/Domain: LDL receptor ligand-binding repeat homology <LDL3>
 F/447-482/Domain: LDL receptor ligand-binding repeat homology <LDL4>
 F/648-682/Domain: LDL receptor ligand-binding repeat homology <LDL5>
 F/684-720/Domain: LDL receptor ligand-binding repeat homology <LDL6>
 F/723-757/Domain: LDL receptor ligand-binding repeat homology <LDL7>
 F/869-1097/Domain: trypsin homology <TRY>

Query Match 18.3%; Score 549; DB 2; Length 1113;
 Best Local Similarity 35.9%; Pred. No. 1.6e-22;

Matches 137; Conservative 60; Mismatches 135; Indels 50; Gaps 18;

QY 199 CPKIAVRCGVVDCUKLSDELGCVRPDMDKSLIKIYSGSSHOWLPICSSNMNDSYS 254
 DB 735 CVRPDMDCGVVDCSDSDSEWCVTLKNGNSSSLTVHKSKEHN---VCADGWRETLIS 791
 QY 255 EKTQQLGFSARHTTEVVAHRD-----FANSFSLRYNSTIOESL-HRSECPQRYIS 306
 DB 792 QLACKQMLGSPSYTKLIPGOEGOWRLYPNWEVL--NSTLQELVLYRRHSCPSRSIS 849
 QY 307 LQGS--HCGLR---AMTRIVGALASDSKMPQWSLHFGTT-HICGGLTIDAQWVLTAA 360
 DB 850 LLCSKQDCGRRPAAARKKILIGRTSRGRMPWCQSLQSESGHICGCVLAKKAVLYVA 909
 QY 361 HCFVTRKVLG-----WKVYAGTSLHQLPEA---LPEASIAEIIINSNTDEEDDYDIA 411
 DB 910 HCF-----EGNEDADVWVVFVGINNLDHPSGWMQRFYKTLILHRYBRAVVDYDIS 961
 QY 412 LMRSLKPLTSLAHHPACLPNHGQFTSLNETCWTGTFKTRTDKTSPPFLREVQVNLID 470
 DB 962 VVELSDDDINETSYPVCLPSPEBYLEPDTCYITGWGHH-----GNKMPFLQEBEYVLI 1017
 QY 471 DFKCNDLYVD-SYLPFRMKAGDLRGGRDSCQGSGLPVCBRO--NNRWYLAGVTSWGT 528
 DB 1018 PLEQCSY--EDMKTIITRMICAGIESGVDSCHMDSGGPLVCEPFGQMTLFGITSNGS 1075
 QY 529 GCGQR-NKPGYVTKVTEVLPMTI 549
 DB 1076 VCFSKVLGPGVYVSNVSYFVWMI 1097

QY 227 DKSL-----LKITYSGSH-----OMLPICSSNMNDSEKTCQ 259
 Db 366 DVLGLTCVPDSGHRISPEFIRLVDGKKEGRVEFVNGWGTTICDGDWTDKHAAYCR 425
 QY 260 QLGFEASARTEVAHRDPANSFILRYNST-----IOESLRSECPQRYISL 307
 Db 426 QLGKGPAPARTMAY--FGEKGPIHMDNVKCTGNKALADCVKDDICRHNCRHSEDAGV 483
 QY 308 QC-----SHCGLRAM---TGRIVGALLADSKRPQVYSLHFGTTH-- 344
 Db 484 ICDYLEKKASSGNGKEMLSGGGLRLHRRQRIIGNNNSLRGAMPQASILRSANHDG 543
 QY 345 --ICGGTIDQWVLTAAHCFVTRKXLEGMKVYAGTSLN-----QLDEA 389
 Db 544 RLTCGATLLSCWVLTAAHCF-----KRIKNNRSYAVAVGDIYHTLVPEEFQOE 592
 QY 390 ASIAEIIINSYVTEBEDDYDIALMLSLKP-----LTLNASHIPACLP--HGQTFSLNET 442
 Db 593 IGVQOIVIHRYRPRRSYDIALVRLQPGRCARLSTHVLPAQLRREPRQKTAAN-- 650
 QY 443 CWTFEGKTRTDTKTSFPLREVOYNLIDFKKNDYLVYDYLPRMCAQDLRGR--D 500
 Db 651 CHITGWTGTRAYSR--LQAAVPLPKRCKR--YKGLFTGRMLCAGNLQEDNRVD 705
 QY 501 SCQGDSSGGLVCEQ--NNRWYLAGVTSKGTGCGQRKPGVYTKTEVLPWIS 551
 Db 706 SCQGDSSGGLVCEQ--NNRWYLAGVTSKGTGCGQRKPGVYTKTEVLPWIS 757

RESULT 12

KFHUI
 coagulation factor Xla (EC 3.4.21.27) precursor (validated) - human
 N/Alternate names: antihemophilic factor C; plasma thromboplastin antecedent
 C/Species: Homo sapiens (man)
 C/Date: 13-Aug-1986 #sequence revision 26-May-1994 #text_change 05-Oct-2004
 C/Accession: A27431; A00920; A37940
 R/Aaskai, R.; Davie, E.W.; Chung, D.W.
 Biochemistry 26, 7221-7228, 1987
 A/Title: Organization of the gene for human factor XI.
 A/Reference number: A27431; MUID:88107663; PMID:2827746
 A/Accession: A27431
 A/Molecule type: DNA
 A/Residues: 1-625 <ASA>
 A/Cross-references: UNIPROT:P03951; UNIPARC:UPI000000DB87; GB:M18295
 A/Note: the sequence shown follows the authors' translation
 R/Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
 Biochemistry 25, 2417-2424, 1986
 A/Title: Amino acid sequence of human factor XI, a blood coagulation factor with four
 A/Reference number: A00920; MUID:86243360; PMID:3636155
 A/Accession: A00920
 A/Molecule type: mRNA
 A/Residues: 1-625 <FTU>
 A/Cross-references: UNIPARC:UPI000000DB87; GB:M1142; NID:g182832; PIDN:AAAS2487.1; PID:
 Biochemistry 30, 2056-2060, 1991
 R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 A/Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
 A/Reference number: A37940; MUID:91152017; PMID:198667
 A/Accession: A37940
 A/Molecule type: protein
 A/Residues: 28-33,35-49,'X',51-55,'X',57-63,70-75,'X',77-79,107-109,'X',111-112,132-139;
 1280-282,'X',284,285-297,313-316,'X',318-319,320-326,'X',328-330,'X',347-349,373,'X',375
 A/Cross-references: UNIPARC:UPI0000172824; UNIPARC:UPI0000172825; UNIPARC:UPI0000172826;
 B2B; UNIPARC:UPI000017282C; UNIPARC:UPI000017282D; UNIPARC:UPI000017282E; UNIPARC:UPI000
 10000172834; UNIPARC:UPI0000172835; UNIPARC:UPI0000172836; UNIPARC:UPI0000172837; UNIPAR
 C/Comment: The proenzyme consists of two identical chains linked by one or more disulfid
 he active site, and a heavy chain, which associates with high molecular weight (HMW) kin
 C/Genetics:
 A/Gene: GDB:F11
 A/Cross-references: GDB:119891; OMIM:264900
 A/Map position: 4q35-q435
 A/Intron: 19/1; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 52
 C/Function:
 A/Description: catalyzes the proteolytic activation of coagulation factor IX

A/Pathway: blood coagulation intrinsic pathway
 C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology
 C/Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
 F/1-18/Domain: signal sequence, duplication; glycoprotein; hemophilia C; homodimer; hydro
 F/19-387/Product: coagulation factor Xla heavy chain #status experimental <HCH>
 F/119-108/Domain: apple repeat <AP1>
 F/109-198/Domain: apple repeat <AP2>
 F/199-288/Domain: apple repeat <AP3>
 F/290-379/Domain: apple repeat <AP4>
 F/388-625/Product: coagulation factor Xla light chain #status experimental <LCH>
 F/388-618/Domain: trypsin homology <TRY>
 F/20-103,514-581,571-599/Disulfide bonds: #status predicted
 F/29/Disulfide bonds: interchain #status experimental
 F/46-76,50-56,110-193,136-165,140-146,200-283,226-255,230-236,291-374,317-346,321-327,380
 F/90,126,353,450/Binding site: carbohydrate (asn) (covalent) #status predicted
 F/339/Disulfide bonds: interchain #status predicted
 F/387-388/Cleavage site: Arg-116 (coagulation factor XIa) #status experimental
 F/431,480,575/Active site: His, Asp, Ser #status predicted
 F/491/Binding site: carbohydrate (asn) (covalent) #status experimental

Query Match 17.7%; Score 532; DB 1; Length 625;
 Best Local Similarity 33.2%; Pred. No. 7,2e-22;
 Matches 131; Conservative 62; Mismatches 127; Indels 74; Gaps 17;

QY 212 CLKSDKELG--CYRPMDKSLKITYSGSHQWLP-C-SNMND-----SYSE 255
 Db 255 CLKTSBGLPSTIKKSKALSGSLSCRSIFVCHSSYHDTDFGLBELDVAASH 314
 QY 256 KTCQOLGFESARHTEVAHRDPANSFILRYNSTIOESLRSECPQRYISLQCSHGLR 315
 Db 315 EACQKL-----CTNAVRCQF--FTYTPAASCNBQ--KGKC---YKLKSLNGSPTK 358
 QY 316 AMTG-----RYGGLALADSKRPQVYSLHFGTTH-----HIC 346
 Db 359 ILHRGGISGTYLRLCKNDECTTKIKPRIYIGGTAASVGEPMQVTLH--TTSPTORLQ 416
 QY 347 GGLTIDQWVLTAAHCF--VTRKVLSEGMKYAGTSLNLCQPEAS--IAEIIINSYV 402
 Db 417 GGSIIIGNWMTLTAHCFVCSPIIL--RYSGILNOSKEIKEDTSFPGVDEIIHDQY 473
 QY 403 DEBDYDIALMLSKPLTSLAHPACLPVHGQTFSLNETCMTGTFGKTRTDTKTSFPL 462
 Db 474 MAESGYDIALKLETNYVYTSQRPICPSKGDNRNVIYTDQWVGMG--YRLKLRDQIQT 532
 QY 463 REVQVNLIDFKKNDYLVYDYLTPRMCAQDLRGHSCQGSGLPVCQNNRWYLA 521
 Db 533 OKAKIPLVTNECCOR--YRGHKITHKMTCAGYREGGCKDCKDPSGLSCKHNEVHTLV 590
 QY 522 GVTSMGTCGCGQRNKPQVYTKTEVLPWISQES 555
 Db 591 GVTSMGTCGCGQRNKPQVYTKTEVLPWISQES 624

RESULT 13

PIBO
 plasmin (EC 3.4.21.7) precursor - bovine
 N/Alternate names: plasminogen
 C/Species: Bos primigenius laurus (cattle)
 C/Date: 30-Sep-1987 #sequence revision 28-Apr-1995 #text_change 09-Jul-2004
 C/Accession: S45046; A25835; I45961; S03736
 R/Berglund, L.; Andersen, M.D.; Petersen, T.E.
 Submitted to the EMBL Data Library, May 1994
 A/Description: Cloning and characterization of the bovine plasminogen cDNA.
 A/Reference number: S45046
 A/Accession: S45046
 A/Molecule type: mRNA
 A/Residues: 1-812 <BER>
 A/Cross-references: UNIPROT:P06668; UNIPARC:UPI0000043BBB; EMBL:X79402; NID:g494962; PID:
 A/Experimental source: liver
 A/Note: it is uncertain whether Met-1 or Met-8 is the initiator
 R/Schaller, U.; Moser, P.W.; Danneberger-Muller, G.A.K.; Rosselet, S.J.; Kamfer, U.; Rick
 Eur. J. Biochem. 149, 267-278, 1985
 A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmi

A:Reference number: A25835; MUID:85203906; PMID:3846532
A:Accession: A25835
A:Molecule type: protein
A:Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
A:Cross-references: UNIPARC:UPI0000172895
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: 145961; MUID:85023311; PMID:6148961
A:Accession: 145961
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 706-743, 'R', 745-812 <MAL>
A:Cross-references: UNIPARC:UPI000016365; GB:K02935; NID:9163551; PIDN:AA30714.1; PID:
R:Brumfiel, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.B.; Lergler, W.; Manneberg, M.;
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: 803735; MUID:81212097; PMID:7238497
A:Accession: 803735
A:Molecule type: protein
A:Residues: 27-83 <BRU>
A:Cross-references: UNIPARC:UPI0000172896
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma;
F:1-36/Domain: signal sequence #status predicted <SIG>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:127-812/Product: plasminogen #status experimental <PRO>
F:127-103/Domain: activation peptide #status experimental <APT>
F:104-583,584-812/Product: plasmin #status experimental <MT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringle homology <KR1>
F:192-268/Domain: kringle homology <KR2>
F:282-359/Domain: kringle homology <KR3>
F:384-461/Domain: kringle homology <KR4>
F:485-564/Domain: kringle homology <KR5>
F:584-812/Domain: plasmin chain B #status experimental <BCB>
F:584-805/Domain: trypsin homology <TRY>
F:56-80,60-68,110-188,131-171,158-183,192-269,195-323,213-252,241-264,282-359,303-342,33
bonds: #status predicted
F:315/Binding site: carbohydrate (Aan) (covalent) #status experimental
F:365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 17.1%; Score 514; DB 1; Length 812;
Best Local Similarity 35.5%; Pred. No. 9e-21;
Matches 122; Conservative 42; Mismatches 126; Indels 54; Gaps 9;

245 CSSNNNDYSRK-----TCQOLGFSAHRTTEVAHRDFANSFILRVNNTIOESLHR 296
485 CMIGCKSTYRGKATTTAGVPCQEWAAQEPHO-----HSITPERNPQSGLER 532
297 SEC-----PSQRYISLQCS--HCGL-----RAMTGRIVGALASD 329
533 NYCRRPDPDVNGPWCYTNMPKRPDYCDVPCQSESSFDCKPVRPKKSGRIVGCVSKP 592
330 SKMPQVSLHFGTTHICGGLTIDAQVLTAAHCFVTBTKVLSGKRVYTAGSNLHQLDEA 389
593 HSMPPQVSLRBSRRCGGLTISPKNVLTAAHC--LDNLTLSFKVILGAN--EKVRE 648
390 ASIARIINSNTDEDDVDIALMRLSKPLTSAIHHPACLPMHQOTSLNETCMTWISFG 449
649 QSVQSLIP--SRIFRPSQADIALKLKSRPAITTKVIRPCLPPTMYMAARRECTITWIG 707
450 KTRRTDDTSPFLREVOVNLIDFKKNDYLVYDSYLTFRPMKAGDLRGGRSGCGDSGSP 509
708 ETQGTFGS--GLTKSAHLPIVENKVCNREYLDGKVKPTLCAHLIGSTSCGDSGSP 765
510 LVCEQNNRMYLAGVTSKGTGCGQRNKPQVYTYTAVTELVPIYTKM 553

Db 766 LVCFEKDYILQGYTSMGLGCRPNKQGVYVSPVYPIEBETM 809
RESULT 14
A57014
Proctasin (EC 3.4.21.-) precursor - human
C:Species: Homo sapiens (man)
C:Dates: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
A:Accession: A57014; A54866
R:Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 270, 13483-13489, 1995
A:Title: Molecular cloning, tissue-specific expression, and cellular localization of huma
A:Reference number: A57014; MUID:95286644; PMID:7768952
A:Accession: A57014
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-343 <RES>
A:Cross-references: UNIPROT:Q16651; UNIPARC:UPI0000046DCB; GB:I41351; NID:9862304; PIDN:
A:Experimental source: prostate
A:Note: parts of this sequence were determined by protein sequencing
R:Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18843-18848, 1994
A:Title: Proctasin is a novel human serine proteinase from seminal fluid. Purification, t
A:Reference number: A54866; MUID:94308140; PMID:8034638
A:Accession: A54866
A:Molecule type: protein
A:Residues: 45-64 <YUN>
A:Cross-references: UNIPARC:UPI0000172AFB
C:Genetics:
A:Gene: GDB:PR858
A:Cross-references: GDB:676446; OMIM:600823
A:Map position: 16p11.2-16p11.2
C:Superfamily: trypsin; trypsin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-44,45-343/Product: proctasin #status predicted <MT>
F:33-44/Domain: proctasin light chain #status predicted <CHL>
F:45-343/Domain: proctasin heavy chain #status predicted <CHH>
F:45-283/Domain: trypsin homology <TRY>
F:323-341/Domain: transmembrane #status predicted <TMH>
F:37-154,70-86,166-244,201-223,234-262/disulfide bonds: #status predicted
F:85,134,238/Active site: His, Asp, Ser #status predicted
F:159/Binding site: carbohydrate (Aan) (covalent) #status experimental

Query Match 17.0%; Score 510.5; DB 1; Length 343;
Best Local Similarity 40.9%; Pred. No. 5.7e-21;
Matches 103; Conservative 42; Mismatches 94; Indels 13; Gaps 5;

312 CGLRPMTRIVGALASDSKMPQVSLHFGTTHICGGLTIDAQVLTAAHCFVTREKVL 371
37 CGV-APQARITGSSAVAGQMPQVSLYEGVHVCQGSIVSEOWVLSAAHCFPSHHK-- 93
372 EGMVYVAGTSLHQLPEASIA---ETIINSYDDEBDVDIALMRLSKPLTSAIHHPA 426
94 BAYEVKLGALHQLDSYSEBAKYSTLKDILPHPSYVQESQGDIALQLSRPTFSRYIRPI 153
429 CLPMHGOTFSINETCMTWISFGTRETDDKTSPL--LREVOVNLIDFKKNDYLVYDS--- 483
154 CLPAAASFPNGHLCTYTGQHVAPSVSLTPKPLQGLEVLRLRETNCGLYNIDAKREE 213
484 --YLTFRPMKAGDLRGGRSGCGDSGSPLVCEQNNRMYLAGVTSKGTGCGQRNKPQVYTK 541
214 PHFQEDWVACAGYEGKDACQDSGSPVSGVGLWYLTGIVSGWACGARNRPGVYTL 273
542 VTEVLPIYTKM 553
274 ASSYASWTQSKV 285

RESULT 15
UC4171
trypcase (EC 3.4.21.59) precursor - rat
N:Alternate names: mast cell trypase

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:05:05 ; Search time 301 Seconds
(without alignments)
1727.106 Million cell updates/sec

Title: US-10-806-370-12
Perfect score: 2999
Sequence: 1 MERUSHGNAAPARTPSAGAS.....TEVLPWYKMESEVRFRKS 562

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2956	98.6	581	1	TMPSD_HUMAN
2	2510.5	83.7	543	1	TMPSD_MOUSE
3	891	29.7	359	2	Q4SPGQ_TETNG
4	777.5	25.9	490	1	TMPS2_MOUSE
5	771.5	25.7	490	2	Q3UKK3_MOUSE
6	770.5	25.7	490	2	Q7TN04_MOUSE
7	769.5	25.7	490	2	Q6P7D7_RAT
8	763.5	25.5	453	2	Q812A6_MOUSE
9	763.5	25.5	475	2	Q3TZ06_MOUSE
10	763.5	25.5	490	2	Q920K3_RAT
11	762.5	25.4	453	1	TMPS3_MOUSE
12	762.5	25.4	453	2	Q2M1G4_MOUSE
13	761.5	25.4	484	2	Q311U4_MACMU
14	759.5	25.3	486	2	Q5PRA6_BRARE
15	755	25.2	484	2	Q311V4_PANPA
16	754	25.1	484	2	Q311V3_PANTR
17	751	25.0	484	2	Q311V5_PANTR
18	744	24.8	484	2	Q311U8_HYLSY
19	740.5	24.7	492	2	Q6GTU7_HUMAN
20	738.5	24.6	492	1	TMPS2_HUMAN
21	738.5	24.6	492	2	Q6T7J3_HUMAN
22	736.5	24.6	538	2	Q5USC7_HUMAN
23	731	24.4	454	1	TMPS3_HUMAN
24	713.5	23.8	722	2	Q6NUP5_XENLA
25	712.5	23.8	767	1	Q9DGR2_XENLA
26	697.5	23.3	437	2	TMPS4_HUMAN
27	696	23.2	445	2	Q8CJ17_RAT
28	689.5	23.0	435	2	Q5RDX7_PONPY
29	684.5	22.8	435	1	TMPS4_MOUSE
30	684	22.8	388	2	Q4RRR7_TETNG
31	668.5	22.3	371	2	Q8CJ16_RAT

32	664	22.1	455	1	TMPS5_MOUSE	Q9ER04_mus musculus
33	659	22.0	455	2	Q8CDR0_MOUSE	Q8CDR0_mus musculus
34	644.5	21.5	457	1	TMPS5_HUMAN	Q9H3A3_homo sapiens
35	616.5	20.6	445	2	Q3U0U6_MOUSE	Q3U0U6_mus musculus
36	611.5	20.4	417	1	HEPS_HUMAN	P05981_homo sapiens
37	609.5	20.3	417	2	Q5R5E8_PONPY	Q5R5E8_pongo pygma
38	608.5	20.3	436	1	HEPS_MOUSE	Q35453_mus musculus
39	600	20.0	326	2	Q7Z280_BRARE	Q7Z280_brachydanio
40	592.5	19.8	416	1	HEPS_RAT	Q05511_rattus norv
41	580.5	19.4	730	2	Q4RH70_TETNG	Q4RH70_tetradon n
42	578.5	19.3	572	1	TMPS7_MOUSE	Q8B1K6_mus musculus
43	578	19.3	799	1	TMPS6_MOUSE	Q9D410_mus musculus
44	578	19.3	799	2	Q6PF94_MOUSE	Q6PF94_mus musculus
45	578	19.3	811	2	Q3KN88_MOUSE	Q3KN88_mus musculus

ALIGNMENTS

RESULT 1
TMPSD_HUMAN STANDARD; PRT; 581 AA.
ID TMPSD_HUMAN Q99YE2; Q86YM4; Q960Y8; Q99YE1;
AC Q99YE2; Q86YM4; Q960Y8; Q99YE1;
DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 15-MAR-2005, sequence version 2.
DT 07-MAR-2006, entry version 29.
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
DE protease) (Membrane-type mosaic serine protease).
GN Name=TMPS13; Synonyms=MSP, TMPS11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=21167393; PubMed=11267681; DOI=10.1016/S0167-4781(01)00184-1;
RC TISSUE=Lung;
RA Kim D.R., Sharmir S., Inoue M., Kido H.,
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209 (2001).
[2]
RA Nucleotide sequence [MRNA] (ISOFORM 2).
RP Pak T.J., Park W.J.,
RC PubMed=14702039; DOI=10.1038/ng1285;
RX Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niimiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kinata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T., Kusano Y.,
RA Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita H., Imose N.,
RA Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yochaisho K., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shintzu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunuma K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Matanabe M., Konatsu T.,
RA Mitsuhashi-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.",
RL Nat. Genet. 36:40-45(2004).
RN [4].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG Human chromosome 11 international sequencing consortium,
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1 SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
CC protein (Potential).
CC -1 ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=MSP1, large form, Membrane-type;
CC IsoId=Q9BYE2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9BYE2-2; Sequence=VSP_013103, VSP_013104;
CC Note=No experimental confirmation available;
CC Name=3; Synonyms=MSPS, Small form, VSP_013102;
CC IsoId=Q9BYE2-3; Sequence=VSP_013099, VSP_013102;
CC Name=4;
CC IsoId=Q9BYE2-4; Sequence=VSP_013100, VSP_013101;
CC Note=No experimental confirmation available;
CC -1 TISSUE SPECIFICITY: Isoform 1 and isoform 3 are predominantly
CC expressed in lung, placenta, pancreas, and prostate. Isoform 3 is
CC weekly expressed in testis and peripheral blood lymphocytes.
CC -1 POLYMORPHISM: The repeat A-S-P-A-[GLQR] is polymorphic and the
CC number of copies varies between 12 to 14.
CC -1 SIMILARITY: Belongs to the peptidase S1 family.
CC -1 SIMILARITY: Contains 1 LDI-receptor class A domain.
CC -1 SIMILARITY: Contains 1 SKCR domain.
CC -1 CAUTION: Ref.2 has referred to this protein as TMPRSS6.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License

FT TRANSMEM 161 181
FT FT
FT TOPO_DOM 182 581
FT REPEAT 9 13
FT REPEAT 14 18
FT REPEAT 19 23
FT REPEAT 24 28
FT REPEAT 29 33
FT REPEAT 34 38
FT REPEAT 39 43
FT REPEAT 44 48
FT REPEAT 49 53
FT REPEAT 54 58
FT REPEAT 59 63
FT REPEAT 64 68
FT REPEAT 69 73
FT REPEAT 74 78
FT REPEAT 79 83
FT REPEAT 84 88
FT DOMAIN 190 320
FT DOMAIN 199 221
FT DOMAIN 321 554
FT REGION 9 88
FT REGION 14 68
FT COMBIAS 9 89
FT ACT_SITE 361 361
FT ACT_SITE 409 409
FT ACT_SITE 506 506
FT CARBOHYD 250 250
FT CARBOHYD 287 287
FT CARBOHYD 400 400
FT CARBOHYD 440 440
FT DISULFID 245 309
FT DISULFID 258 312
FT DISULFID 346 362
FT DISULFID 443 512
FT DISULFID 475 491
FT DISULFID 502 530
FT VASAPPLIC 146 181
FT FT
FT VASAPPLIC 423 486
FT FT
FT FT
FT VASAPPLIC 555 581
FT FT
FT VASAPPLIC 555 581
FT FT
FT VASAPPLIC 559 581
FT FT
FT CONFLICT 187 187
FT CONFLICT 192 192
FT CONFLICT 201 201
FT CONFLICT 206 206
FT CONFLICT 259 259
FT CONFLICT 298 298
FT CONFLICT 496 496
FT SEQUENCE 581 AA; 62680 MW; A4935CCCF31D29EE CRC64;
SQ

Query Match 98.6%; Score 2956; DB 1; Length 581;
Best Local Similarity 99.8%; Pred. NO. 3.3e-155;
Matches 553; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MERDSHGASPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQAS 60
DB 1 MERDSHGASPARTSPASGASPAQASPAQASPAQASPAQASPAQASPAQAS 60
Oy 61 PAGTPPGASPGASPAQASPARASPALASLSRSSSGSSGRSASASVTTSPTRYVYLVAT 120

```

Db 61 PACGPPGASPGASPAQASPARASPALASLSRSSSSGSSSSASASVTTSPRYLVAT 120
Qy 121 PVGAVPTSSPARSPATRAATRESPTSLPKPTWEGGKQPLGCVULLATLVSLITL 180
Db 121 PVGAVPTSSPARSPATRAATRESPTSLPKPTWEGGKQPLGCVULLATLVSLITL 180
Qy 181 FQFMGHTGIRYKEQRESCPKIAVACDGVDCCKLSDELGCYRPFMDKSLTKYSGSSHQ 240
Db 181 FQFMGHTGIRYKEQRESCPKIAVACDGVDCCKLSDELGCYRPFMDKSLTKYSGSSHQ 240
Qy 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANSPSILKYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANSPSILKYNSTIOESLHRSCEP 300
Qy 301 SQRYSISLQCSHCGLAMTGRIVGALASDSKMPQVSLHFGTTHCGSTLLDAQVVLTA 360
Db 301 SQRYSISLQCSHCGLAMTGRIVGALASDSKMPQVSLHFGTTHCGSTLLDAQVVLTA 360
Qy 361 HCFVTRKRVLEGMKVYAGTSLNLHLPASIAETIIINSNTDEEDYDIALMRLSKPLT 420
Db 361 HCFVTRKRVLEGMKVYAGTSLNLHLPASIAETIIINSNTDEEDYDIALMRLSKPLT 420
Qy 421 LSAHIHPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 480
Db 421 LSAHIHPACLPMHGQTFSLNETCWTGFGKTRTDDKTSPLREYQVNLIDFKKNDYLV 480
Qy 481 YDSYLTTPMNCAGDILRGDRSCQDGGPPLVCEONNRYLAGVTSWGCQCGKRGVYT 540
Db 481 YDSYLTTPMNCAGDILRGDRSCQDGGPPLVCEONNRYLAGVTSWGCQCGKRGVYT 540
Qy 541 KYTEVLPWYSKME 554
Db 541 KYTEVLPWYSKME 554

```

RESULT 2

TMPSD MOUSE STANDARD; PRT: 543 AA.

```

AC Q50405; Q8CFE0; Q9IV08;
DT 15-MAR-2005, integrated into UniProtKB/Swiss-Prot.
DT 07-MAR-2006, entry version 16.
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
DE protease) (Membrane-type mosaic serine protease).
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=B6/EGFP, and Trophectoderm stem cell;
RC TISSUE=Mammary tumor, and Trophectoderm stem cell;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Aleaschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hallyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
CC protein (Potential).
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -1- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL: BC010843; AAH10843.1; -, mRNA.
DR EMBL: BC042878; AAH42878.1; -, mRNA.
DR EMBL: BC085323; AAH85323.1; -, mRNA.
DR HSSP: P00760; 1EXZ.
DR MEROPS: S01.087; -.
DR MGI: MGI:2682935; Tmpres13.
DR InterPro: IPR002172; LDL_rcpt_A.
DR InterPro: IPR001254; Peptidase_S1_S6.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR001190; Src_rcpt.
DR Pfam: PF00057; Ldl_recept_a; 1.
DR Pfam: PF00530; SRCR; 1.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00202; SR; 1.
DR SMART: SM00202; Tryp_Spc; 1.
DR PROSITE: PS01209; LDLRA_1; FALSE_NEG.
DR PROSITE: PS00668; LDLRA_2; FALSE_NEG.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS50287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR GlycoProtex: GlycoProtex; Membrane; Protease; Repeat; Serine protease;
KW Signal-anchor; Transmembrane.
KW CHAIN 1 543
FT 1 TOPO_DOM 1 143
FT 2 TRANSMM 144 164
FT 3 REPEAT 165 543
FT 4 REPEAT 14 17
FT 5 REPEAT 18 22
FT 6 REPEAT 23 27
FT 7 REPEAT 28 31
FT 8 REPEAT 32 36
FT 9 REPEAT 37 40
FT 10 REPEAT 41 45
FT 11 REPEAT 46 49
FT 12 REPEAT 50 54
FT 13 REPEAT 55 59
FT 14 REPEAT 60 64
FT 15 REPEAT 65 69
FT 16 DOMAIN 180 202
FT 17 DOMAIN 199 301
FT 18 DOMAIN 302 535
FT 19 REGION 14 49
FT 20 REGION 18 69
FT 21 COMBIDS 11 72
FT 22 ACT_SITE 342 342
FT 23 ACT_SITE 390 390
FT 24 ACT_SITE 487 487
FT 25 CARBOHYD 231 231
FT 26 CARBOHYD 268 268
FT 27 CARBOHYD 381 381
FT 28 CARBOHYD 421 421
FT 29 DISULFD 226 290
FT 30 DISULFD 239 293
FT 31 DISULFD 327 343
FT 32 DISULFD 424 493

```

FT DISULFID 456 472 By similarity.
 FT DISULFID 483 511 By similarity.
 FT CONFLICT 281 S -> P (in Ref. 1; AAH42878).
 FT CONFLICT 475 475 D -> G (in Ref. 1; AAH85323).
 FT CONFLICT 530 530 I -> T (in Ref. 1; AAH85323).
 SQ SEQUENCE 543 AA; 59806 MW; 61026D04A0FCB2D5 CRC64;
 Query Match 83.7%; Score 2510.5; DB 1; Length 543;
 Best Local Similarity 83.3%; Pred. No. 1.7e-131;
 Matches 468; Conservative 40; Mismatches 35; Indels 19; Gaps 4;
 QY 1 MERBSHGASAPRPPASAPASPAQSPATPPGRASPAQASPAQASPAQAS 60
 DB 1 MDROSHRNSPAPRT-----PPOASPARTSPARA-----PPOASPARTPP-----QAS 42
 QY 61 PATPPPGASPORASPAQASPAQASPALASLSRSSGSSSSARSASVTTSPTRVLYVAT 120
 DB 43 PARPP- QASPARAPPPASPARASPARAPPSRSSGSSSSARSASTTSPTRVLYVAT 101
 QY 121 PVGAVPIRSSPARASAPATRATRESPGTSLPKPTWRBQKQDPLIGCVLLIALVVSITL 180
 DB 102 PVGAVPIRASPASAPATRATRESPGTSLPKPTWRBQKQDPLIGCVLLIALVVSITL 161
 QY 181 FQPMQATGIRYKQRESCPKAARCDGVNDCILKSDCLGCTRPMDKSLIKIYSSSHQ 240
 DB 162 FVPRGHGIRYKKEPLBSCPIHAVACDVNCKMSDELGCVRPMDKSLIKYSSSGE 221
 QY 241 WLPISSWMDNSYSKTCQQLGFBSAHRHTTEVAHDPANFSFILYNGSTIOESLHRSCEP 300
 DB 222 WLPVSSWMDNTDSKTCQQLGFBSAHRHTTEVAHDPANFSFILYNGSTIOESLHRSCEP 281
 QY 301 SORVLSLQCHSGCLPAMTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQVLTAA 360
 DB 282 SRRVYSLQCHSGCLPAMTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQVLTAA 341
 QY 361 HCFPTREKTLFGMVYVAGTSLHQLPRAASIAETIINSNTDEEDVDIALMLSKLT 420
 DB 342 HCFPTREKTLFGMVYVAGTSLHQLPRAASIAETIINSNTDEEDVDIALMLSKLT 401
 QY 421 LSAHHPACLPWNGQTFSLNETCMTTGFKTRTETDNTSPFLREVOVNLIDFKKNDYLV 480
 DB 402 LSAHHPACLPWNGQTFSLNETCMTTGFKTRTETDNTSPFLREVOVNLIDFKKNDYLV 461
 QY 481 YDSYLTTPRMWAGDLRGDRSCQGDGSGPLVCEONRWYLAGVTSWGTGCGQRNPGYTT 540
 DB 462 YDSYLTTPRMWAGDLRGDRSCQGDGSGPLVCEONRWYLAGVTSWGTGCGQRNPGYTT 521
 QY 541 KYTEVLPWYISKMSSEVRFKRS 562
 DB 522 KYTEVLPWYISKMSSEVRFKRS 543
 RESULT 3
 Q4SPG0_TETNG PRELIMINARY; PRT; 359 AA.
 AC Q4SPG0;
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Chromosome 16, SCF41537, whole genome shotgun sequence. (Fragment).
 GN ORFNames=GSTENG001484901;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 CC Tetraodontidae; Tetraodontidae; Tetraodon.
 CX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RP PubMed=15496914; DOI=10.1038/nature03025;
 RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Strange-Thomann N.,
 Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Daeilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anhouard V., Jubin C., Castell J., Katinka M., Vacherie B.,
 RA Blomont C., Skallit Z., Cattolico L., Poullain J., De Berardis V.,
 RA Craud C., Duprat S., Broctier P., Couancan J.-P., Gouy J.,
 RA Parra G., Lardier G., Chappie C., McKernan K.J., McGowan P., Bosak S.,
 RA Keller M., Volff J.-N., Guigo R., Zody M.C., Mestrov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Landel V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RT Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RP Genoscope, Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC -----
 CC EMBL: CAAB01014537, CAP97472.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005044; F:scavenger receptor activity; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR001190; Strc_rcpt.
 DR Pfam: PF00089; Typein_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; Tryp_SPC_1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 FT NON_TER 1
 FT TER 359
 SQ SEQUENCE 359 AA; 38943 MW; 4C14083C7823B37 CRC64;
 Query Match 29.7%; Score 891; DB 2; Length 359;
 Best Local Similarity 44.2%; Pred. No. 1.1e-41;
 Matches 159; Conservative 70; Mismatches 121; Indels 10; Gaps 4;
 QY 202 HAVCDGVNDCILKSDCLGCVRPMDKSLIKIYSSSHQWLPISSWMDNSYSKTCQQL 261
 DB 1 NATHCDGVNDCILKSDCLGCVRPMDKSLIKIYSSSHQWLPISSWMDNSYSKTCQQL 59
 QY 262 GFESAHRTTEVAHDPANFSFILRYNSTIOESLH-----RSECPGRIYSLQCHGCLRA 316
 DB 60 GFNFVYATNPSTSQ--PKSSPTLTINSRSSPYLQGRVNVSSSCPGQGVVALQCLDCCGR 117
 QY 317 MTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQVLTAAHCFVTEKVL--EGW 374
 DB 118 STSRIGNVAKLQWPMQMTLHFRGSHVCGILISPDVLTAAHCFESNKLALAINW 177
 QY 375 KYVAGTSLHQLPRAASIAETIINSNTDEEDVDIALMLSKLTLSAHPACLPWNG 434
 DB 178 EYSGVESLIDLPLPYKTKRILLSELYNSDNDVDALLKLAAPVFPDNPQACLPBRD 237
 QY 435 QTFSLNETCMTTGFKTRTETDNTSPFLREVOVNLIDFKKNDYLVDSYLTTPRMWAG 494
 DB 238 QILAPGTCMTTGFKTRTETDNTSPFLREVOVNLIDFKKNDYLVDSYLTTPRMWAG 297
 QY 495 LRGRDSCQGDGSGPLVCEONRWYLAGVTSWGTGCGQRNPGYTTKYTEVLPWYISKMS 554
 DB 298 LRGRDSCQGDGSGPLVCEONRWYLAGVTSWGTGCGQRNPGYTTKYTEVLPWYISKMS 357
 RESULT 4
 TMPS2_MOUSE STANDARD; PRT; 490 AA.

AC Q99108; Q99UK4; Q99Y82;
DT 11-JAN-2001, integrated into UniProtKB/Swiss-Prot.
DT 11-JAN-2001, sequence version 2.
DT 07-MAR-2006, entry version 52.
DE Transmembrane protease, serine 2 (EC 3.4.21.-) (Epitheliasin) (Plasmic
DE transmembrane protein X) [Contains: Transmembrane protease, serine 2
DE non-catalytic chain; Transmembrane protease, serine 2 catalytic
DE chain].
GN Name=Tmpres2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND TISSUE SPECIFICITY.
RC STRAIN=BALB/c;
RX MEDLINE=21104370, Pubmed=11165526;
RA DOI=10.1096/9896.2000.9999.9999;:AID-PATH743>3.0.CO;2-T;
RA Vazara M.H., Potvati K.S., Keilokumpu S., Kyllonen A.P., Vihko P.T.;
RT "Expression of transmembrane serine protease TMPRSS2 in mouse and
RT human tissues.";
RL J. Pathol. 193:134-140(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Han J., Kim S.;
RT "Putative transmembrane protease X.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=BALB/c;
RX PubMed=10683448; DOI=10.1016/S0014-5793(00)01196-0;
RA Jacquelinet E.J., Rao N.V., Rao G.V., Hoidal J.R.;
RT "Cloning, genomic organization, chromosomal assignment and expression
RT of a novel mosaic serine proteinase: epitheliasin.";
RL FEBS Lett. 468:95-100(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=129; TISSUE=Mammary gland;
RX MEDLINE=22386257, Pubmed=12477932, DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Stauber R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Alschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marnuska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepieton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smailus D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
CC cleavage and secreted (by similarity).
CC -1- TISSUE SPECIFICITY: Expressed mainly in prostate and kidney.
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
DR EMBL, AF199362; AAF97867.1; -, mRNA.
DR EMBL, AF243500; AAF64186.1; -, mRNA.

[illegible]

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komno H., Aachii J., Fukuda S.,
 RA Aikawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kaibawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Flisbachmann W., Gaasterland T., Gissi C., King B., Kochava H.,
 RA Kuehl P., Lewis S., Matcuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Strubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guentrich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeert P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszeh-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsunki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
 RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RN Genome Res. 10:1617-1630(2000).
 RN [7]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aikawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Komno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RN Genome Res. 10:1757-1771(2000).
 RN [8]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Placenta;
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
 RA Kawai J., Kojima M., Komno H., Murata M., Nakamura M., Ninomiya N.,
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watanabe A.,
 RA Muramatsu M., Hayashizaki Y.;
 RT Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL, AK146047; BAE26858.1; -; mRNA.
 DR MGI, MGI:1354381; Tmp082.
 DR GO, GO:0016021; C: integral to membrane; RCA.
 DR InterPro, IPR002172; LDI_rcpt_A.
 DR InterPro, IPR001254; Peptidase_S1_S6.
 DR InterPro, IPR001314; Peptidase_S1A.
 DR InterPro, IPR001190; Strc_rcpt.
 DR Pfam, PF00057; Ldi_recept_a; 1.
 DR Pfam, PF00089; Trypsin_1.
 DR PRINTS, PR00722; CHYMOTRYPSIN.
 DR SMART, SM00192; LDLa; 1.
 DR SMART, SM00202; SR; 1.
 DR SMART, SM00202; TYP_SPC; 1.
 DR PROSITE, PS01209; LDLa_1; 1.
 DR PROSITE, PS0068; LDLa_2; 1.
 DR PROSITE, PS0287; SRCL_2; 1.

DR PROSITE, PS0240; TRYPSIN_DOM; 1.
 DR PROSITE, PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE, PS00135; TRYPSIN_SER; 1.
 KM Hydrolase; Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 490 AA; 5349 MW; 54651AF288E76EBA C6C64;
 Query Match 25.7%; Score 771.5; DB 2; Length 490;
 Best Local Similarity 31.5%; Pred. No. 6,66-35;
 Matches 174; Conservative 87; Mismatches 200; Indels 91; Gaps 13;
 QY 27 AGTPGGRASPQASPAQASPAQSPAGTPGGRASPQASPAQATPPGGRASPQASPAQASPA---R 83
 DB 5 SGSPPG-IGPCYEHNGYSEHICPP-----RPAPAPNGYNLPYQYPSVPVQ 51
 QY 84 ASPPLASISRSSSSSSARSASVTSPTRYLVATPVGAVPVIRSSPARSAPATRYRE 143
 DB 52 YAPRTTQASTSVTHTHPKSSGALCTSSKSLCLALALGV----- 93
 QY 144 SPGTSLPFTWRREGQKQLPLIGCVLLLTALVSVSLILFQPMQHTGIRYKQRES--CPK 201
 DB 94 -----LTGAANAVNLNRPVDSNCSGSEMGSGTGCS 127
 QY 202 HAVRCGVVDCKLKSLDELGVCFPMWKSLLKIYSGSSHQWLPICSSNNNDYSKTCQQL 261
 DB 128 SLMWCDGVAAHCPNGEDENRCVRLYGQSFILQVYSSQRAKAWPVVCCDDWSESGRAACKDM 187
 QY 262 GPESAHRTTE-VAHKDFPANSFILRYNS---TIQESLHRECECPQRIYISQCHGCLR 316
 DB 188 GYKNFFYSSQGIIPQSGATFSMKLVNVSNGVNLKYLHSDSCSSRMVSLRCIEGVR 247
 QY 317 M--TCRIYVGLALASDPKPMQVSLFPTTHCGTLLDAQWLVLAHCFVTRKVLBG- 373
 DB 248 VKQSGRIYVGLANASBGMFPMQVSLHVGAVVAGSGSITPEMIVPAHCV---EELPS 303
 QY 374 --WKYVACT-----SNLQLEPAASIAIINSNTYDEBDYDIALMRLSKPLTUSA 423
 DB 304 RYWTAFACILQSLQSLMFYSGRQ-----VEKVISHPNYSKYKNDIALMKLQTLAIFND 357
 QY 424 HIHPACLFMGQTSLSNCTGTRGKTRFDDKTSPLREYQVNLIDPKKNDVLVYDS 483
 DB 358 LVKPVCLPSPGMMLDDECEWISGQATYE-KGKTSVDVLANAMVLIPEKNSKRYLYNN 416
 QY 484 YLTPRMGAGDLRGGRDSQGDGSGPLVCEONRMYLVAGTVSWGTCGORNKPYVTKYT 543
 DB 417 LITPMITCAGFLQGSVDSQGGSGPLVTLKNGIMWLIQDTSWGGCAKALRPGYGVNT 476
 QY 544 EVLPWYSKMS 555
 DB 477 VFTDWIKQMR 488
 RESULT 6
 OTTN04 MOUSE PRELIMINARY; PRT; 490 AA.
 AC OTTN04.
 DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2003, sequence version 1.
 DT 21-FEB-2006, entry version 16.
 DE Transmembrane protease, serine 2.
 GN Name=Tmp082;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=129; TISSUE=Mammary tumor, Brcal-/f1;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shmeman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleon M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RM [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=129; TISSUE=Mammary tumor. Brcal-/E1;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: BC054348; AAH54348.1; -; mRNA.
 CC HSSP: P00760; 1EZX.
 CC Ensemble: ENSMUSG00000000385; Mus musculus.
 CC MGI: MGI:1354381; Tmpres2.
 CC GO: GO:0016021; C:integral to membrane; RCA.
 DR InterPro: IPR0012172; Ldb_rcpt_A.
 DR InterPro: IPR001254; Peptidase_S1_S6.
 DR InterPro: IPR001190; Srrc_rcpt.
 DR Pfam: PF00057; Ldl_recept_a; 1.
 DR Pfam: PF00530; SRCR; 1.
 DR Pfam: PF00089; Tryptan; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00192; LDLA; 1.
 DR SMART: SM00202; SR; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS01209; LDLRA_1; 1.
 DR PROSITE: PS00068; LDLRA_2; 1.
 DR PROSITE: PS00287; SRCR_2; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
 SO SEQUENCE 490 AA; 53526 MW; 546508028417655A CRC64;
 Query Match 25.7%; Score 770.5; DB 2; Length 490;
 Best Local Similarity 31.5%; Pred. No. 7.5e-35;
 Matches 174; Conservative 87; Mismatches 200; Indels 91; Gaps 13;
 Oy 27 AGTPGRASPAQASPAQASPACTPPGRASPAQASPACTPPGRASPAQASPAQASPA---R 83
 Db 5 SGPSPG-IGPCVENGYSSEHICP-----PPVAPNPNYNYLPAYQYSPVQ 51
 Oy 84 ASPALASISRSSSGSSSARSASAVTSPRYLVATVGAVPVISPSPARASPARATRR 143
 Db 52 YAPRTTQASTSVHTHKSGALCTSKSKSLCALALGV----- 93
 Oy 144 SPGTSLPKFRREQOKPLIGCVLLILALVSLILFQFQNGHTGIRYKQRS--CPK 201
 Db 94 -----LGAAVAALVLRFPDSDNSTSEMGSSGTCIS 127
 Oy 202 HAVKCDGVVDCXKSLDELGCYRPFMDKSLTKYSSSSHQMLPTGSSMNDSTSEKTCQOL 261
 Db 128 SSIWCDGVGAHCPNGDENRCVRLYGQSPILQYSSQKRAWYPCDDWSSESGRAACDM 187
 Oy 262 GFESAHRTTE-VAHRDFANSFSLRLYNS---TIGESLHRSCEPQRIYISLQCHSGIRA 316
 Db 188 GYKNNFYSSQGIPODSGATSPKLVNVSNGVNDLYKLYHSDSCSSRMVYSIRCIEGVRS 247

Oy 317 M--TGRIVGALASDPSKMPQVSLHFGTTHICGGLIDAQWLTAHCFVTRKYLEG- 373
 Db 248 VKRGSRIYGINASGDMPMQVSLHVGQVHVGGSIIIPEMIVTAHCV-----EELPSSP 303
 Oy 374 --WKVYAGT-----SNLHQLPEASIAETIINSNTYDEEDYDALMRSLKPLTUSA 423
 Db 304 RYWTFAFATLRQSLMFYSRHQ-----VEKVISHPNDSKTKNNDIAMKLTQPLAEND 357
 Oy 424 HIHPCLPMHQTFSLNETCWITGPKTRERDDKTSPPFLREVQVNLIDPKCNDLYVDS 483
 Db 358 LKPVCLPFPGMMLDLDECMISGATYE--KKTSDVLAAMVPLIPSKNSKYITNN 416
 Oy 484 YLTPRMQCALRGGRDSCQDGGPLVCEONNRWYLAGVTSWGTGCGQNRKPGVYTYT 543
 Db 417 LITPAMICAGFLQGSVDS CQDSSGFLVTLKNGITWMLIGDTSWGSCKALRPGYGVNT 476
 Oy 544 EYLPWYISKMS 555
 Db 477 VFTDWIYQOMRA 488
 RESULT 7
 Q6P7D7_RAT PRELIMINARY; PRT; 490 AA.
 ID Q6P7D7;
 AC Q6P7D7;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 21-FEB-2006, entry version 18.
 DE Transmembrane protease, serine 2.
 GN Name=Tmpres2;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.D., Collins F.S., Wagner L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L.H., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleon M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RM [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: BC061712; AAH61712.1; -; mRNA.
 CC HSSP: P00761; 1AKS.
 CC Ensemble: ENSMUSG00000001976; Rattus norvegicus.
 DR GO: GO:0016021; C:integral to membrane; IEA.

OY 494 DLRGGRDSCQSDSGPLVCEQNNRWYLAGVTSWGTGCGCRNPGVYTKTEVLPMWYSKM 553
 DB 388 YLKGSDVSCQSDSGPLVCEQNNRWYLAGVTSWGTGCGCRNPGVYTKTEVLPMWYSKM 447
 OY 554 ESEVR 558
 DB 448 ERDLK 452

RESULT 9
 ID Q37206_MOUSE PRELIMINARY; PRT: 475 AA.
 AC Q37206;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 21-FEB-2006, entry version 7.
 DE Adult inner ear cDNA, RIKEN full-length enriched library,
 clone:R930041P06 product:titinmembrane protease, serine 3, full insert
 sequence.
 GN Name=Tmpres3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 CC Muridae; Murinae; Murinae; Mus.
 NCBI TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 Bajic M.V., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilm L.G., Aldrich V., Allen J.E.,
 Ambesi-Impombato R., Apweiler R., Aurrelija R.N., Bailey T.L.,
 Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 Fletcher C.F., Funkhouser T., Guleras T.R., Gajdardzic J., Gariboldi M.,
 Georgi-Jacob M., Goldberg I., Grimbergen J., Green R.B.,
 Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keisio J., Kikuchi H.,
 Kikuchi H., Kohler S., Kuhlmann S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lazarev L.F., Lazarevic D., Lipovich L., Liu J.,
 Liu J., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 McBryden E., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mutsaers J., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Nilsen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Ohtsuka V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Roest B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
 Schaubach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,
 Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 Sperling S., Stupka E., Sugita K., Sultana R., Takanaka Y., Taki K.,
 Tamada K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen B., Vercato R., Wei C.L., Yang K.,
 Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide W., Bult C.,
 Wilmshurst S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 Willemsen C., Mettlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura N.,
 Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
 Tegami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,

RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 Bajic M.V., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilm L.G., Aldrich V., Allen J.E.,
 Ambesi-Impombato R., Apweiler R., Aurrelija R.N., Bailey T.L.,
 Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 Fletcher C.F., Funkhouser T., Guleras T.R., Gajdardzic J., Gariboldi M.,
 Georgi-Jacob M., Goldberg I., Grimbergen J., Green R.B.,
 Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keisio J., Kikuchi H.,
 Kikuchi H., Kohler S., Kuhlmann S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lazarev L.F., Lazarevic D., Lipovich L., Liu J.,
 Liu J., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 McBryden E., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mutsaers J., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Ohtsuka V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Roest B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
 Schaubach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,
 Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 Sperling S., Stupka E., Sugita K., Sultana R., Takanaka Y., Taki K.,
 Tamada K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen B., Vercato R., Wei C.L., Yang K.,
 Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide W., Bult C.,
 Wilmshurst S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 Willemsen C., Mettlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura N.,
 Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
 Tegami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,
 Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 Bajic M.V., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilm L.G., Aldrich V., Allen J.E.,
 Ambesi-Impombato R., Apweiler R., Aurrelija R.N., Bailey T.L.,
 Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 Fletcher C.F., Funkhouser T., Guleras T.R., Gajdardzic J., Gariboldi M.,
 Georgi-Jacob M., Goldberg I., Grimbergen J., Green R.B.,
 Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keisio J., Kikuchi H.,
 Kikuchi H., Kohler S., Kuhlmann S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lazarev L.F., Lazarevic D., Lipovich L., Liu J.,
 Liu J., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 McBryden E., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mutsaers J., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Ohtsuka V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Roest B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
 Schaubach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,
 Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 Sperling S., Stupka E., Sugita K., Sultana R., Takanaka Y., Taki K.,
 Tamada K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen B., Vercato R., Wei C.L., Yang K.,
 Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide W., Bult C.,
 Wilmshurst S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 Willemsen C., Mettlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura N.,
 Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
 Tegami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,
 Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 Bajic M.V., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 Davis M.J., Wilm L.G., Aldrich V., Allen J.E.,
 Ambesi-Impombato R., Apweiler R., Aurrelija R.N., Bailey T.L.,
 Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 Chiu K.P., Choudhury V., Christoffels A., Clutterbuck D.R.,
 Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 Fletcher C.F., Funkhouser T., Guleras T.R., Gajdardzic J., Gariboldi M.,
 Georgi-Jacob M., Goldberg I., Grimbergen J., Green R.B.,
 Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
 Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keisio J., Kikuchi H.,
 Kikuchi H., Kohler S., Kuhlmann S.P., Kruger A., Kummerfeld S.K.,
 Kurochkin I.V., Lazarev L.F., Lazarevic D., Lipovich L., Liu J.,
 Liu J., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 McBryden E., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 Mutsaers J., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 Okazaki Y., Ohtsuka V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 Roest B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
 Schaubach C., Sekiguchi K., Sempole C.A., Sene S., Sessa L., Sheng Y.,
 Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 Sperling S., Stupka E., Sugita K., Sultana R., Takanaka Y., Taki K.,
 Tamada K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 Ueda H.R., van Nimwegen B., Vercato R., Wei C.L., Yang K.,
 Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide W., Bult C.,
 Wilmshurst S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 Willemsen C., Mettlick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura N.,
 Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
 Tegami M., Waki K., Watanabe A., Okamura-Ohno Y., Suzuki H., Kawai J.,
 Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563 (2005).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=inner ear;
 RX PubMed=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

[illegible]

Db 177 HLRVLDLEQFOGDF--VSINHLSDDKVATLHHSVYRRECTGHHVTLKCSACGR-- 233

Qy 318 TG---RIVGALASDSKPMQVSLHFGTTHICGTLIDAQVLTAAHCFVYREKVLGW 314

Db 233 TGVSPRIVGCGNMSSLTQMPQVSLQFGCVHLGGSIITPLVIYAAACVVDLYHP--KSW 290

Qy 375 KYVAGTNSNHLQPEAASIAE-IINSNATDEEDYDIALMLSKPLSLAHIPALCPMH 433

Db 291 TVQVGLVSLMDSVPVSHLEVKEIIVHSKPKPRGLGNDIALMLKSEPLTFDETIQICLPNS 350

Qy 434 GQFSLNTECMWTGKGTRETDKTSPLRBYQVNLIDFKCNDYLVYDSVLTFRMNCAG 493

Db 351 EENFPDGLKCMVSGKATEDGGD-ASPVLNHAAPVLISNKCINRDVYGGIISPMLCAG 409

Qy 494 DLKGRDSCQSGSGGPLVCEQNNRWYTLAVTSSWGTGCGQRNKPQVYTKVTEVLPWISKM 553

Db 410 YLKGVDSCQSGSGGPLVQGERLMLKVLGANSFGIGCAEVNKKPVYTRITSFLLWHIQL 466

Qy 554 ESEVR 558

Db 470 ERDLK 474

RESULT 10

Q920K3_RAT PRELIMINARY; PRT; 490 AA.

AC Q920K3; 0920K3; 01-DEC-2001, integrated into UniProtKB/TrEMBL.

DT 01-DEC-2001, sequence version 1.

DT 21-FEB-2006, entry version 24.

DE TMPRSS2.

GN Name=TmpRSS2; Synonyms=TMPRSS2;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

LN RATTUS NORVEGICUS.

RA Tsuzuki S.

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane protein (By similarity).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL; AB073550; BAB70683.1; -; mRNA.

DR HSSP; P00760; 1EZK.

DR MEROPS; S01.247; -.

DR Ensembl; ENSRNOG0000001976; Rattus norvegicus.

DR RGD; 620766; TmpRSS2.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005044; F:scavenger receptor activity; IEA.

DR GO; GO:0004552; F:serine-type endopeptidase activity; IEA.

DR GO; GO:0006058; P:proteolysis; IEA.

DR InterPro; IPR002172; LDL_rcpt_A.

DR InterPro; IPR001254; Peptidase_S1_S6.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR001190; Src_rcpt.

DR Pfam; PF00057; Ildl_rcpt_a; 1.

DR Pfam; PF00089; Trypsin; 1.

DR PRINTS; PRO0722; CHYMOTRYPSIN.

DR SMART; SM00192; LDLa; 1.

DR SMART; SM00202; SR; 1.

DR SMART; SM00202; TRYP_Spc; 1.

DR PROSITE; PS50068; LDLRA_2; 1.

DR PROSITE; PS50287; SRCR_2; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

DR Hydrolase; Membrane, protease; Serine protease; Transmembrane.

QO SSQUNCE 490 AA; 53519 MW; 286691551CA0409A CRC64;


```

OY 211 DCKLDELGVCPDWDKSLIKIYSGSSHOWLPICSSNMWDSYSEKTCQQLGF-----BSA 266
DB 97 DCKNAEDRYRCVRSGGQALQVFTAA--WRTWCSDMKSHYAKIACAQGFPEYSSD 154
OY 267 HRTTEVAHRDPANSPSILRYNSTIOE---SLH-----RSECPQRYISLQCSHGLPAM 317
DB 155 HLRVDALAEQFGDPF--VSINHLSDDKYVTLAHSVVMRECGTSGHVVTLKCSACGTF-- 210
OY 318 TG---RIVGALASDSKMPQVSLHFGTTHICGGLIDAQVYLTAAHCFVTRREKVLGWM 374
DB 211 TGYSPRIVGGMSSLTQMPQVSLQFGYHLCGGSVITPLMTVTAACVYDLYHP--KSW 268
OY 375 KYVAGTSLHQLPEASIAE--IINSNTDEEDVDIALMRSLKPLTSAHHPACLPMM 433
DB 269 TVQGVLSLMDSPVPSHLVEKIYHSKYPKRLGNDIALMLKSEPLTDEDTIQTCLPNS 328
OY 434 GQTSINETCMTITGFKTRTDDKTSPELREVQVNLIDFKCNDLYVDSYLTPEPMGAG 493
DB 329 EENPFDGKLCWTSGMGATEDGD--ASPVLNHAAVPLISNKCINHRDVGIIISPMCLAG 387
OY 494 DLRGGRSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGQRNKPQVYTKVTEVLPWYISKM 553
DB 388 YLKGGVDSQGDGSGPLVCEQERRMLKVLGATSFGLICAEVKNKPGVYTRITSFLLMHIQOL 447
OY 554 ESEVR 558
DB 448 ERDLK 452

```

RESULT 12

Q2M1G4_MOUSE PRELIMINARY; PRT; 453 AA.

AC Q2M1G4;
 DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.
 DT 21-FEB-2006, sequence version 1.
 DE Transmembrane protease, serine 3.
 GN Name=Impress3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=PCR rescued clones;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Alechni S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bork S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=PCR rescued clones;
 RG NIH MGC Project;
 RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC112375; AAI12376.1; -, mRNA.
KW Protease; Transmembrane.
SQ SEQUENCE 453 AA; 49492 MW; 1ABCBF10AF6E1EF6 CRC64;

Query Match      25.4%; Score 762.5; DB 2; Length 453;
Best Local Similarity 38.6%; Pred. No. 1.9e-34;
Matches 164; Conservative 68; Mismatches 150; Indels 43; Gaps 13;

OY 161 LPL-----IGCVLLALVAVSLILFQF---WQHGIRYKQRESCPKHAAVCDQV 210
DB 44 LPLKFPPIVIGITALLIALIGLIFHDCGKTRCHSSFR-----CIELTRACDVS 96
OY 211 DCKLDELGVCPDWDKSLIKIYSGSSHOWLPICSSNMWDSYSEKTCQQLGF-----BSA 266
DB 97 DCKNAEDRYRCVRSGGQALQVFTAA--WRTWCSDMKSHYAKIACAQGFPEYSSD 154
OY 267 HRTTEVAHRDPANSPSILRYNSTIOE---SLH-----RSECPQRYISLQCSHGLPAM 317
DB 155 HLRVDALAEQFGDPF--VSINHLSDDKYVTLAHSVVMRECGTSGHVVTLKCSACGTF-- 210
OY 318 TG---RIVGALASDSKMPQVSLHFGTTHICGGLIDAQVYLTAAHCFVTRREKVLGWM 374
DB 211 TGYSPRIVGGMSSLTQMPQVSLQFGYHLCGGSVITPLMTVTAACVYDLYHP--KSW 268
OY 375 KYVAGTSLHQLPEASIAE--IINSNTDEEDVDIALMRSLKPLTSAHHPACLPMM 433
DB 269 TVQGVLSLMDSPVPSHLVEKIYHSKYPKRLGNDIALMLKSEPLTDEDTIQTCLPNS 328
OY 434 GQTSINETCMTITGFKTRTDDKTSPELREVQVNLIDFKCNDLYVDSYLTPEPMGAG 493
DB 329 EENPFDGKLCWTSGMGATEDGD--ASPVLNHAAVPLISNKCINHRDVGIIISPMCLAG 387
OY 494 DLRGGRSCQGDGSGPLVCEQNNRWYLAGVTSWGTGCGQRNKPQVYTKVTEVLPWYISKM 553
DB 388 YLKGGVDSQGDGSGPLVCEQERRMLKVLGATSFGLICAEVKNKPGVYTRITSFLLMHIQOL 447
OY 554 ESEVR 558
DB 448 ERDLK 452

```

RESULT 13

Q311U4_MACMU PRELIMINARY; PRT; 484 AA.

AC Q311U4;
 DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
 DT 08-NOV-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Transmembrane protease serine 2 (Fragment).
 GN Name=IMPRESS2;
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 CX NCBI_TaxID=9544;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Clark N.L., Swanson W.J.;
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 EMBL; DOI:50506; AA82295.1; -, Genomic DNA.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0005044; F:scavenger receptor activity; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR002172; LDL_rcpt_A.
 DR InterPro; IPR001254; Peptidase_S1_S6.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001190; Srcr_rcpt.
 DR Pfam; PF00057; Ldl_recept_a; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLa; 1.
 DR SMART; SM00202; SR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01209; LDLa_1; 1.
 DR PROSITE; PS00068; LDLa_2; 1.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Protease; Transmembrane.
 FT NON_TER 1
 FT NON_TER 484
 SQ SEQUENCE 484 AA; 53082 MW; 09E3BED8739792B0 CRC64;

Query Match 25.3%; Score 761.5; DB 2; Length 484;

Best Local Similarity 33.3%; Pred. No. 2,3e-34;
 Matches 185; Conservative 81; Mismatches 188; Indels 101; Gaps 19;

QY 28 GTPPGR-----SPAQASPAQSPAGTPPG--RASPAQSPAGTPPGRASPGASPA 77
 DB 1 GSPGVGVYENHGVQENPYPAQPTVA---PNVEVHPAOVYPPV-----QYTPR 50
 QY 78 QASPARASPALASLSRSSGSSSARASVTTSPRYVLVRAVPYGAVPINSSPARASPA 137
 DB 51 VLTTH-SNPVAVCORQKPSPG---TVCTSKTKKALCVMTLGAV----- 89
 QY 138 TRARRESGTSIPKRTWRGQKQLPLIGCVLLTLVLSIILFQFQ-----GHTGIR 191
 DB 90 -----LVGA-----ALAAGL-----WKFGSKGSDSGLS 114
 QY 192 YKEQRESQPKHAIVCDGVVCKLXSDDELGCVRFDWDKSLIKYSGSSHQMLPTCSSMND 251
 DB 115 C-DSGCTCISSNMCDGVSHCPNGEDENRCVRLVGNPNIILQVYSQGRSMHPVCRDDNE 173
 QY 252 SYSEKTCQQLPESAHRTTE-VARHDFANSFSLRY---NSTIGSLHRSF-CPSQRYIS 306
 DB 174 NYAABAACDMGVKNSFYSOGIVDSGATSPFKLNTSAGNDIYKCLYHSDACSKAVVS 233
 QY 307 LQCHGGR---AMTGRIVGALASDSKMPPOVSLHFTTHICGTLIDAGMVLTAACF 363
 DB 234 LRCTACGSRNLSRSLRSGRIAGQNALGAMPPOVSLHVNIVCGSILTPETVTAACV 293
 QY 364 FVTEKVLG---WKVYAGTSLNHL-PEAASIAEIIINSNYTDEDDYDIALMRLSKP 418
 DB 294 -----EKLPSWQMTAFVGTILRSSMFYKGRHVEKXVSHNRYNSKTONIDALMKLTP 349
 QY 419 LTLSAHHPACLPNMGQTFSLNETCWITGFKTRETDTKTSPPLEVOVNLIDFKCNDY 478
 DB 350 LTFEIVVAVPCLPMPGMLEPEQHWMISGWGATGE-KKSTDVLAANAAPPILPERCNK 408
 QY 479 LVVDSYLLPRMWCAGDLRGDRSCGDSGGLVCEONNRWYLAIGTSGTCGGRNKGV 538
 DB 409 YVYGLITPAKICAGFLQTVDSGDSGGLVTLKNDVMMILGITSNGSGCAANRRGV 468
 QY 539 YTKVTEVLPWYISKM 553
 DB 469 YGNVTVFTDWIYRQM 483

RESULT 14
 QSPRA6 BRARB PRELIMINARY; PRT; 486 AA.
 AC QSPRA6
 DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
 DT 04-JAN-2005, sequence version 1.

DT 21-FEB-2006, entry version 10.
 DE Zgc:101791.
 GN Oryzanes-zgc:101791;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Olfactory epithelium;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Alechuth S.F., Zeeberg B., Bluetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Murnane K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tomihata S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Heltun B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Olfactory epithelium;
 RL Director MGC Project; to the EMBL/GenBank/DBJ databases.
 RU Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type II membrane
 CC protein (By similarity).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL; BC086738; AA86738.1; - mRNA.
 DR Ensemble; ENSDARG0000045889; Danio rerio.
 DR ZFIN; ZDB-GENE-041212-48; Zgc:101791.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0005044; F:scavenger receptor activity; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis; IEA.
 DR InterPro; IPR002172; LDL_rcpt_A.
 DR InterPro; IPR001254; Peptidase_S1_S6.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001190; Srcr_rcpt.
 DR Pfam; PF00057; Ldl_recept_a; 1.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLa; 1.
 DR SMART; SM00202; Tryp_Spc; 1.
 DR PROSITE; PS01209; LDLa_1; 1.
 DR PROSITE; PS00068; LDLa_2; 1.
 DR PROSITE; PS50287; SRCR_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KM Hydrolyase; Membrane; Protease; Serine protease; Transmembrane.
 SQ SEQUENCE 486 AA; 53213 MW; 8618788C5F5A13C CRC64;

Query Match 25.3%; Score 759.5; DB 2; Length 486;
 Best Local Similarity 34.1%; Pred. No. 3e-34;
 Matches 190; Conservative 77; Mismatches 177; Indels 113; Gaps 19;

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model1

Run on: September 16, 2006, 03:13:35 ; Search time 50 Seconds

(without alignments)
983.844 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASPARPPSAGAS.....TEVLPWYKMSSEVRFRKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /EMC_Celerra_SIDS3/prodata/2/1aa/5 COMB.pap: *
2: /EMC_Celerra_SIDS3/prodata/2/1aa/6 COMB.pap: *
3: /EMC_Celerra_SIDS3/prodata/2/1aa/7 COMB.pap: *
4: /EMC_Celerra_SIDS3/prodata/2/1aa/H COMB.pap: *
5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCUS COMB.pap: *
6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE COMB.pap: *
7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfile1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2999	100.0	562	2	US-09-879-792-12
2	2291.5	76.4	477	2	US-10-177-661-2
3	2108	70.3	446	2	US-10-177-661-4
4	969.5	32.3	418	2	US-10-177-661-6
5	741.5	24.7	492	2	US-09-759-143-932
6	741.5	24.7	492	2	US-10-012-896-932
7	741.5	24.7	492	2	US-10-144-678A-932
8	740.5	24.7	492	2	US-09-342-749-2
9	740.5	24.7	492	2	US-09-691-840-2
10	740.5	24.7	510	2	US-09-949-016-11074
11	736.5	24.6	492	2	US-09-685-166A-895
12	736.5	24.6	492	2	US-09-879-792-14
13	736.5	24.6	492	2	US-09-679-426-895
14	736.5	24.6	492	2	US-09-759-143-895
15	736.5	24.6	492	2	US-10-012-896-895
16	736.5	24.6	492	2	US-10-144-678A-895
17	736.5	24.6	521	2	US-09-949-016-11081
18	736.5	24.6	521	2	US-09-949-016-11082
19	736.5	24.6	521	2	US-09-949-016-11083
20	729.5	24.3	453	2	US-09-999-833A-69
21	729.5	24.3	453	2	US-10-020-445A-69
22	729.5	24.3	453	2	US-09-978-188-69
23	729.5	24.3	453	2	US-10-017-085A-69
24	729.5	24.3	453	2	US-10-145-129A-69
25	729.5	24.3	453	3	US-10-013-929A-69
26	729.5	24.3	453	3	US-10-013-917A-69

27	727.5	24.3	393	2	US-09-759-143-934	Sequence 934, App
28	727.5	24.3	393	3	US-10-012-896-934	Sequence 934, App
29	727.5	24.3	393	2	US-10-144-678A-934	Sequence 934, App
30	726	24.2	454	2	US-09-518-046-2	Sequence 2, Appli
31	726	24.2	454	2	US-09-650-371-2	Sequence 2, Appli
32	704	23.5	423	2	US-09-656-002-2	Sequence 2, Appli
33	697.5	23.3	435	2	US-09-607-745-2	Sequence 2, Appli
34	697.5	23.3	435	3	US-10-030-688-2	Sequence 2, Appli
35	697.5	23.3	437	2	US-09-851-588-8	Sequence 8, Appli
36	696.5	23.2	435	2	US-09-008-271A-6	Sequence 6, Appli
37	696.5	23.2	435	2	US-09-968-415-6	Sequence 6, Appli
38	695.5	23.2	406	2	US-09-851-588-6	Sequence 6, Appli
39	694	23.1	432	2	US-10-012-231A-275	Sequence 275, App
40	694	23.1	432	2	US-10-015-389A-275	Sequence 275, App
41	694	23.1	432	2	US-10-006-768A-275	Sequence 275, App
42	694	23.1	432	2	US-10-015-671A-275	Sequence 275, App
43	694	23.1	432	2	US-10-015-393A-275	Sequence 275, App
44	694	23.1	432	2	US-10-011-833A-275	Sequence 275, App
45	694	23.1	432	2	US-10-006-041A-275	Sequence 275, App

ALIGNMENTS

```
RESULT 1
US-09-879-792-12
Sequence 12, Application US/09879792
Patent No. 6734006
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
FILE REFERENCE: 02973.00035
CURRENT FILING DATE: 2001-06-13
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-09-879-792-12
Query Match 100.0%; Score 2999; DB 2; Length 562;
Best Local Similarity 100.0%; Pred. No. 1.9e-204;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MERSHGNASPARPPSAGASPAQASPACTPPGRASPAQASPAQASPACTPPGRASPAQAS 60
1 MERSHGNASPARPPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
61 PAGTPPGASGRASPAQASPARASPALASLSRSSSGRSSARASVTTSTPTRYLVAT 120
61 PAGTPPGASGRASPAQASPARASPALASLSRSSSGRSSARASVTTSTPTRYLVAT 120
121 EVGAVPIRSSPARAPATRATRESPTGSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
121 EVGAVPIRSSPARAPATRATRESPTGSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
121 EVGAVPIRSSPARAPATRATRESPTGSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
121 EVGAVPIRSSPARAPATRATRESPTGSLPKFTWREGQKPLIGCVLLIALVSLIIL 180
181 FQFQGHGHTGTYKQKQSCPGHAYRCGVNDCKLKSDELGVRFPMWDSLTKIYSSSHQ 240
181 FQFQGHGHTGTYKQKQSCPGHAYRCGVNDCKLKSDELGVRFPMWDSLTKIYSSSHQ 240
241 WLPICSSNMNDYSEKTCQQLGFESAHRTTEVAHRDFANSFSLIKYNSTIOESLHRSCEP 300
```

Db	241	WLPICSSMWNDSYSEKTCQQLQJFESHRRTTEVAHNDFAISFSLRNRNSTIIQBSLHSECP	300
Qy	301	SQRYSLQCSHCGLPAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWLTAA	360
Db	301	SQRYSLQCSHCGLPAMTGRIVGALASDSKMPQVSLHFGTTHICGTLIDAQWLTAA	360
Qy	361	HCFPTREKVLGEMRYVAGTSMHLQPEAASTAEIININNYDEEDDYDIALMRLSKPLT	420
Db	361	HCFPTREKVLGEMRYVAGTSMHLQPEAASTAEIININNYDEEDDYDIALMRLSKPLT	420
Qy	421	LSAHIHPACLPMHGQTFSLNTECWTGFGKTRERTDKTSPLREVOVNLIDFKKCDYLV	480
Db	421	LSAHIHPACLPMHGQTFSLNTECWTGFGKTRERTDKTSPLREVOVNLIDFKKCDYLV	480
Qy	481	YDSYITPRMCMAGDLRGGRDSCQGSQGGPLVCBQNNRWYLAGYTSWGTGCGQANRKGVT	540
Db	481	YDSYITPRMCMAGDLRGGRDSCQGSQGGPLVCBQNNRWYLAGYTSWGTGCGQANRKGVT	540
Qy	541	KYTEVLPMIYSQMESEVFRKS	562
Db	541	KYTEVLPMIYSQMESEVFRKS	562

```

RESULT 2
US-10-177-661-2
/ Sequence 2, Application US/10177661
/ Patent No. 6794173
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, Dirk M.
/ APPLICANT: Vitrca, G. Duke
/ TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
/ FILE REFERENCE: 3256-A
/ CURRENT APPLICATION NUMBER: US/10/177,661
/ CURRENT FILING DATE: 2002-06-20
/ PRIOR APPLICATION NUMBER: US 60/299,606
/ PRIOR FILING DATE: 2001-06-20
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 477
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-177-661-2

```

[illegible]

QY	DB	QY	DB
480	395	540	455
VDSYLLTPRMKCGDLRGGRDSCQDSGPLYCEONRMTYAGTSMGTGCGQNRKGVY	VDSYLLTPRMKCGDLRGGRDSCQDSGPLYCEONRMTYAGTSMGTGCGQNRKGVY	TKTETVLPMTYAGTSMSEVAFRRKS	TKVTBLVPMYISQMSSEVAFRRKS
539	454	562	477

```

RESULT 3
US-10-177-661-4
; Sequence 4, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Vlica, G. Duke
; TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
; FILE REFERENCE: 3256-A
; CURRENT APPLICATION NUMBER: US/10/177,661
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/299,606
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-661-4

```

Query Match	70.3%	Score 2108	DB 2108	Length 446
Best Local Similarity	88.7%	Pred. No. 1.9e-141		
Matches 393	Conservative 4	Mismatches 12	Indels 34	Gaps 2
QY	120	TPYGAVPINSSSPARSAIPATRAIRBSBPGTSLPKFTWRBSQOKPLIGCYLLIALIVSLIT	179	
DB	38	SPLTSSISSSPASPAVALLL--GTSLPKFTWRBQOKPLIGCYLLIALIVSLIT	94	
QY	180	LPGFMQGHGIRIKREORRESCPGHAYRCQGVYDCKLKSPDLGCVAFPMDDKSLTKIYSGSSH	239	
DB	95	LPGFMQGHGIRIKREORRESCPGHAYRCQGVYDCKLKSPDLGCVAFPMDDKSLTKIYSGSSH	154	
QY	240	QMLPICSSNMNDYSYSKTKCQQLGFESAHRTTEVAHRDFANSFSLRYNSTIQQSLHRSEC	299	
DB	155	QMLPICSSNMNDYSYSKTKCQQLGFESAHRTTEVAHRDFANSFSLRYNSTIQQSLHRSEC	183	
QY	300	PSORYSLSCSHCGRAMTGRIVGALALASDKMPQVSLHGTTHICGTTILDAQWVLT	359	
DB	184	PSORYSLSCSHCGRAMTGRIVGALALASDKMPQVSLHGTTHICGTTILDAQWVLT	243	
QY	360	AHCFVYTRKVLBGMKRVAGTSLHQLPEASIAEIIINSYTDDEDDYDIALMRLSKPL	419	
DB	244	AHCFVYTRKVLBGMKRVAGTSLHQLPEASIAEIIINSYTDDEDDYDIALMRLSKPL	303	
QY	420	TLSAHHHPACLPMHQGTFSLNETCMITPGKTRTDDTSPFLREVOYNLIDPKKNDYL	479	
DB	304	TLSAHHHPACLPMHQGTFSLNETCMITPGKTRTDDTSPFLREVOYNLIDPKKNDYL	363	
QY	480	VYDSYLLTPRMKCGADLRGRDSCQGDSSGAPLVCEBONNRWYLAGVTSMTGCGQGNKPGVY	539	
DB	364	VYDSYLLTPRMKCGADLRGRDSCQGDSSGAPLVCEBONNRWYLAGVTSMTGCGQGNKPGVY	423	
QY	540	TKYTEVLPMIYSRMESEVAFRKS 562		
DB	424	TKYTEVLPMIYSRMESEVAFRKS 446		

```

RESULT 4
US-10-177-661-6
; Sequence 6, Application US/10177661
; Patent No. 6794173
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.

```

APPLICANT: Vitca, G. Duke
TITLE OF INVENTION: DENDRITIC CELL TRANSMEMBRANE SERINE PROTEASE
FILE REFERENCE: 3256-A
CURRENT APPLICATION NUMBER: US/10/177,661
CURRENT FILING DATE: 2002-06-20
PRIORITY APPLICATION NUMBER: US 60/299,606
PRIORITY FILING DATE: 2001-06-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 418
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Consensus sequence
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)..(418)
OTHER INFORMATION: Xaa = unknown
US-10-177-661-6

Query Match 32.3%; Score 969.5; DB 2; Length 418;
Best Local Similarity 50.2%; Pred. No. 8.2e-61;
Matches 208; Conservative 27; Mismatches 130; Indels 49; Gaps 12;

OY 198 SCFPAHRCGVDVCKLSDCLG-----VRPDMKSLIKIYSGSSHOMLPICS 246
DB 1 SCIFXKXWCDVXDCPEXGDEXKCVXXXXXGPPXXVRLYGDXXLLQVYSSXXXXXKPVCS 60
OY 247 SNNDSYSEKTCQOLGFESAHRTTE--VAHRDFANSFSLR-----YNSTIOES 293
DB 61 DNMESYXXKXQMGKXXSAXYXSKXXKXGANSFKLVNXPXNLLXHXYSXIXXX 120
OY 294 LHRSE--CPGQRYISLQCS--HCGLR---AMTRIVGALASDKPMQVSLHFGT-T 343
DB 121 LXRSSXXKCPGKXVVSLOCSXQDCGVRINAXXMTSRIVGXKXASXGKMPQVSLQXXGV 180
OY 344 HICGGLIDAQWVLTAAHCFVTRKYLEGKVVAG--TSLHQLPEAA--SIAELIINS 399
DB 181 HLCGSLIXPFWVLTAAHCVGRXXKPLKGMXVFAGILTXSLHXKPKXXKXVEKILHP 240
OY 400 NY-----TDEEDVDIALMRSLKPLTLAAHHPACLPMHGOTFSINTECMTGFKTRE 453
DB 241 NYXXXXXXKXKXNDIALMRSLKPLTLTXDYIOPVCLNPGOXLXPGTTXIXGKATXE 300
OY 454 TDDKTSPLREVOVNLIDFKCNDYLVYDSYLTTPRMCAQDLRGGRDSCQDGGPLVCE 513
DB 301 XXGTSPLVLEAXVPLIDNKKXNSYXYDXNITPRMTCAGVLEGGVDSQDGGPLVCE 360
OY 514 ---QNNRWYLAGVTSWGTGC-GQNRKGVYTKTEVLPWYISKMESEVRRKS 562
DB 361 XXXXQNNRWMLXGXTSMGXGCAKAKPKGVYTXVYXFLXWYSQKXEXRFRKS 414

RESULT 5
US-09-759-143-932
Sequence 932, Application US/09759143
Patent No. 6800746

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stoik, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hegler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: DIAGNOSIS OF PROSTATE CANCER
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-932

Query Match 24.7%; Score 741.5; DB 2; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

OY 62 AGTPRGRA-----SPGRASPAOAS-----PARASPALASLSRSSGSSRSARS 104
DB 5 SGSPPALGPPYENHGYOPENDYPAQPTVPTVYVHPQYYP----- 46
OY 105 ASVTSPTRVYLVATPVGAVPIRSSPARSAPATRATRESPTSLPKFTWREGQKOLPL- 163
DB 47 ---SPVQYAPRVLTQASNPVCTQPKSPSGVCTSKT-----KKALCIT 88
OY 164 --ICGVLLIALVLSLILFQWQ-----GHTGIRYKQRESCPKAIVRCGVVDCXK 215
DB 89 LTLCTFLVGAALAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPNSMCDGVSHCPG 142
OY 216 SDELGCYRFWDKSLIKIYSGSSHOMLPICSNNNDYSEKTCQOLGFESAHRTTEVAHR 275
DB 143 EDENRCVRLYGSNFILOYSSQKSKMHPVCCDDNNENYGRACADMGKKNFYSSQ-GIV 201
OY 276 DFANSFSLRYSNT-----IOESLHRSE-CPGQRYISLQCSHGLR---AMTRIVGAL 326
DB 202 DSGSTSPKMLNTSAGVNDIYKLYHSDACSSKAVSLRCLACGVNLSRQSRIVGES 261
OY 327 ASDSKPMQVSLHFGTTHICGGLIDAQWVLTAAHCFVTRKYLEG--WKVYAGTSLN 383
DB 262 ALPGAMPQVSLHQNVAHVCSSITPEWITVAHCV---EKPLNPMWHTAFAGILRQ 317
OY 384 HOLPEAA--SIAELIINSNTYDEEDVDIALMRSLKPLTLAAHHPACLPMHGOTFSINE 441
DB 318 SFMFYAGYQYKXISHNVDSTKNDIALMRLOKPLTFNDLVKPVCLPMPGMMLQPEQ 377
OY 442 TCWITGKTRTETDQKTSPLREVOVNLIDFKCNDYLVYDSYLTTPRMCAQDLRGGRD 501
DB 378 LQWISGMKATE-KKTSSEVLNAKAVLLIETORCNSRYVNDLITPMICAGFLQNVDS 436
OY 502 CQDSSGGLVCEQNNRWYLAGVTSWGTGCGQNRKGVYTKTEVLPWYISKMESE 556
DB 437 CQDSSGGLVTSKXNIMWLIDGTSWGSACAAYRPGVYGVNMFDTMIRYMRAD 491

RESULT 6
US-10-012-896-932
Sequence 932, Application US/10012896
Patent No. 6943236

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stoik, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick

APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshinhiro
APPLICANT: Meagher, Medeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-932

Query Match 24.7%; Score 741.5; DB 2; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

62 AGTPPGRA-----SPGRASPAQAS-----PARASPALASLSRSSSGSSARS 104
5 SGSPPAIPYENHGYOPENPYPAQPTVPTVEVHPAQQP----- 46
105 ASVTTSPTRVLYVATVGAVPPIRSSPARSAPATRAATRESPTSLPKFTWREGOKOLPL- 163
47 -----SPVQYAPRVLTOASNPVCTQPKSPSGVCTSKT-----KALCIT 88
164 --ICVLLIALVLSLILFOFWQ-----GHTGIRYKQRESCKPAVRCGVVDCCK 215
89 LTLGTFLVGAALAAAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPSNMCDSVSHCPG 142
216 SDELGCVRPMDKSLKITYSGSSHOWLPICSSMNDSDSEKTCQQLGFGESAHRTTEVAHR 275
143 EDENRCVRLYSGNSFLQYSSSQKSMHPVCCDDMNENGRACRDMGYKNFYSSQ-GIV 201
276 DFANSFSLIRYNSY-----IOESLHRSR-CPQRYISLQCHSGLR---AMTGRIVGAL 326
202 DDSGSTFMKLNITAGVNDIYKQLYHSDACSKAVVSLRCLACGVNLNSSROSRIVGES 261
327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTRKYLEG--WKVYAGTSL 383
262 ALPQAMPQVSLHGVNVHVCSSITPEMIVTAHCV---EKPLNPMHTAFAGLRQ 317
384 HOLPEAA--STAETIINSYTDDEDDVIALMRSLKPLTSAHHPACLPNMGOTFSINE 441
318 SFMFYGAQYQVEKISHPNYDSKTKNNDIALMKLOKPLTFNDLVKPVCLPQPMMLQPEQ 377
442 TCWITGFKTRETDDKTSPLREVOVNLIDFKKNDLVYDSYLTTPRMCMCAGDLRGDS 501
378 LCMWISGMAATEE-KGKTSBVLANAAVLLIETQRNSRYVDNLTTPMIMCAGFLQGVNDS 436
502 CGDSGGLVCEONNRWYLAGVTSWGTGCGGRNKRPGVYTKTEVLPWLYSKMSE 556
437 CGDSGGLVYVSKNNIMWLIGDTSWGSCKAKAYRGGVGNVFTDWTYRQMRAD 491

RESULT 7
US-10-144-678A-932
Sequence 932, Application US/10144678A
Patent No. 7033827
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Mantanabe, Yoshinhiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144,678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 932
LENGTH: 492
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-678A-932

Query Match 24.7%; Score 741.5; DB 3; Length 492;
Best Local Similarity 32.3%; Pred. No. 1.4e-44;
Matches 173; Conservative 85; Mismatches 189; Indels 88; Gaps 16;

62 AGTPPGRA-----SPGRASPAQAS-----PARASPALASLSRSSSGSSARS 104
5 SGSPPAIPYENHGYOPENPYPAQPTVPTVEVHPAQQP----- 46
105 ASVTTSPTRVLYVATVGAVPPIRSSPARSAPATRAATRESPTSLPKFTWREGOKOLPL- 163
47 -----SPVQYAPRVLTOASNPVCTQPKSPSGVCTSKT-----KALCIT 88
164 --ICVLLIALVLSLILFOFWQ-----GHTGIRYKQRESCKPAVRCGVVDCCK 215
89 LTLGTFLVGAALAAAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPSNMCDSVSHCPG 142
216 SDELGCVRPMDKSLKITYSGSSHOWLPICSSMNDSDSEKTCQQLGFGESAHRTTEVAHR 275
143 EDENRCVRLYSGNSFLQYSSSQKSMHPVCCDDMNENGRACRDMGYKNFYSSQ-GIV 201
276 DFANSFSLIRYNSY-----IOESLHRSR-CPQRYISLQCHSGLR---AMTGRIVGAL 326
202 DDSGSTFMKLNITAGVNDIYKQLYHSDACSKAVVSLRCLACGVNLNSSROSRIVGES 261
327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTRKYLEG--WKVYAGTSL 383
262 ALPQAMPQVSLHGVNVHVCSSITPEMIVTAHCV---EKPLNPMHTAFAGLRQ 317
384 HOLPEAA--STAETIINSYTDDEDDVIALMRSLKPLTSAHHPACLPNMGOTFSINE 441
318 SFMFYGAQYQVEKISHPNYDSKTKNNDIALMKLOKPLTFNDLVKPVCLPQPMMLQPEQ 377
442 TCWITGFKTRETDDKTSPLREVOVNLIDFKKNDLVYDSYLTTPRMCMCAGDLRGDS 501
378 LCMWISGMAATEE-KGKTSBVLANAAVLLIETQRNSRYVDNLTTPMIMCAGFLQGVNDS 436
502 CGDSGGLVCEONNRWYLAGVTSWGTGCGGRNKRPGVYTKTEVLPWLYSKMSE 556


```

? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO: 11074
? LENGTH: 510
? TYPE: PRT
? ORGANISM: Human
US-09-949-016-11074

```

Query Match	24.7%;	Score 740.5;	DB 2;	Length 510;
Best Local Similarity	32.3%;	Pred. No. 1.8e-44;		
Matches 173;	Conservative 85;	Mismatches 189;	Indels 88;	Gaps 16;

[illegible]

RESULT 11
US-09-685-165A-895
Sequence 895, Application US/09685165A895
Patent No. 6630305
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlockter, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvyick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel

```

1  APPLICANT:  Wang, Aijun
2  APPLICANT:  Skelky, Vasir A.W.
3  APPLICANT:  Hepler, William
4  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
5  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
6  FILE REFERENCE:  210121.427C21
7  CURRENT APPLICATION NUMBER:  US/09/685,166A
8  CURRENT FILING DATE:  2000-10-10
9  NUMBER OF SEQ ID NOS:  898
10 SOFTWARE:  FastSeq for Windows Version 3.0
11 SEQ ID NO 895
12     LENGTH: 492
13     TYPE:  PRT
14     ORGANISM:  Homo sapiens
15     US-09-685-166A-895

```

Query Match	24.6%;	SCORE	736.5;	DB	2;	Length	492;
Best Local Similarity	32.4%;	Pred. No.	3.2e-44;				
Matches	173;	Conservative	64;	Mismatches	189;	Indels	88;
						Gaps	16

```

Oy 62 AGVHPGKA-----SPPKASPMQAS-----PKAAPALASLBKSSGKSSAKS 100
Db 5 SGSPPAIGPYEENNGYOPEENYPAPQPTVVPVUYEYHNPQOYPP----- 46
Oy 105 ASVTTSPLRVLYVATPAGVAPISAPASAPATRAATBSQTSJLPKFTMEGQGLPL- 163
Db 47 -----SPVPOYAPRVLTQASNPVCTQPKBSGVCYSKST-----KKALCIT 88
Oy 164 --IGCVLLIALVYSLILFQFWO-----GHTSIRYKXQRESCPKAVRCGVDDCYLK 215
Db 89 LTTGTFIVGALAAAGLL-----MKFMGSKCSNGSIBC--DSGGTCINPNMCCDGVSHCPG 142
Oy 216 SDELGCYRFMDKSLKIKYSSGSHQWLPICSSNNNDYSSEKTCOGLGESAHRTEVNAHR 275
Db 143 EDENRCVRLYGPNFLQMTSSQKRSMPHVCCDDMBENHGRACRMYGKKNFYSSQ-GIV 201
Oy 276 DFANSPILRYNST-----IOESLHRSR-CPQRYTSLQCSHCGLR---AMTGRIVGAL 326
Db 202 DDGSGTSFMKLTNTSAGANDVIKKLYHSDAGCSKAVSLRJCLACGVNLNSRQSLVGGES 261
Oy 327 ASDSKMPQVQSLHFGTTHICGGLTIDAQWVLTAAHCFYVTEBKVLG---WKVYAGTSL 383
Db 262 ALPGAMPQVQSLHQAQNVHVCSSITTPMVIYAHCV---EKPINNPMHMTAFEGILRO 317
Oy 384 HOLPEEA--SIATIIINSNYTDEEDDYIALMRLSKPLTSLAHINPACLPMHQQTSPINE 441
Db 318 SFMYEYGAQYQYQXKISHNYDSKTKNNDIALMKOKPLTFPNDIVKVLTPNQMMLQEBQ 377
Oy 442 TCWITFGKTRRETDDKTSPLREYQVNLIDFKKCNVDLYVYSYLTTPRMKACDLRGGRDS 501
Db 378 LCMWISGMGATBE--KGKTSYVLNAKAVLLIEYORCNSRYVYDNLITPAIICAGFLOGANDS 436
Oy 502 CQGSQGGPGLVEONNRWYLAGVTSMGTCGQGRNKRGVYTKTTEVLPMTYSKMS 555
Db 437 CQGSQGGPGLVYSNNINIMWLTGDSITMGSCACAARYRPGVYGVANVWFDMYIRQCKA 490

```

RESULT 12
US-09-879-792-14
Sequence 14, Application US/09879792
Patent No. 6734006
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
APPLICANT: Gedrich, Richard
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
PROTEASE
FILE REFERENCE: 02977.00035
CURRENT APPLICATION NUMBER: US/09/879,792
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: US 60/283,648
 ; PRIOR FILING DATE: 2001-04-16 (Docket No. 6734006 LIO-81-WO)
 ; PRIOR APPLICATION NUMBER: PCT
 ; PRIOR FILING DATE: 2001-06-12
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 492
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-879-792-14

Query Match 24.6%; Score 736.5; DB 2; Length 492;

Best Local Similarity 32.4%; Pred. No. 3.2e-44; Mismatches 189; Indels 88; Gaps 16;

Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;
 QY 62 AGTPPGR-----SPGRASPAQAS-----PARASPALASLRSSSGRSSARS 104
 Db 5 SGSPPAIGPYENHGYQENPYPAQPTVPTVEVHPAQYF----- 46
 QY 105 ASVTPSPRYLVATPVGAVPIRSSPARSAPATRATESPOTSLPKFTWEGOKQLPL- 163
 Db 47 -----SPVQYAPRVLTOASNPVCTQPKSPSGTCTSKT-----KALCIT 88
 QY 164 --IGCVLLIALVLSLILFQFWQ-----GHTGIRYKQRESCEPKAVRCDCGVDCIK 215
 Db 89 LTLGTFVLGALLAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPNCWDGVSHCPGG 142
 QY 216 SDELGCVRFMDKSLIKTYSGSSHOWLPICSSNNDSYSEKTCQOLGFEASHRTTEVAHR 275
 Db 143 EDENRCVRLYGNPNTLQWYSSQKSMHVCQDDMNENYGRAACRMGKKNFYSSQ-GIV 201
 QY 276 DFANSFSLRYNST-----IOESLHRSB-CPSORYISIQCSHGRL--AMTGRIVGAL 326
 Db 202 DDSGSTFMKNTSAGNVDYKQLYHSDACSSKAVVSLRCLACGVNLSSRSQSRIVGES 261
 QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTEKEVLEG--WKVYAGTNTL 383
 Db 262 ALPGAMPQVSLHVNVAHVCSSITTPRMIVTAHCV-----EKPLNNPMHTAFAGILRQ 317
 QY 384 HOLPEAA--SIAETIINSNYDEEDDYDIALMRSLKPTLSAHNIPACLPMHGQFSLNE 441
 Db 318 SFMFYAGYQVQKISHENYDSKTKNDIALMKLOKPLTFNDLVKPVCLPMPGMLOPEQ 377
 QY 442 TCWITGFKETRETDDKTSPLREYQVNLIDPKKNDYLVYDSYLTFRMPCAGDLRGDS 501
 Db 378 LCMISGMWATEE-KGKTSEVINAALVLLIETQRCNSRYVDNLITPAMICAGFLOGNDS 436
 QY 502 CQDSSGGLVCEQNNRWYLAGVTSWGTGCGGRNKGVTYTKTEVLPWITYSKMES 555
 Db 437 CQDSSGGLVTSNNNIMWLIGDTSWGSACAKAYRPGYGVNMFWDWIYRQKA 490

RESULT 13

US-09-679-426-895
Sequence 895; Application US/09679426

Patent No. 6759515

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolck, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yaser A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.427C20
 ; CURRENT APPLICATION NUMBER: US/09/679,426
 ; CURRENT FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 895
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 895
 ; LENGTH: 492
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-679-426-895

Query Match 24.6%; Score 736.5; DB 2; Length 492;

Best Local Similarity 32.4%; Pred. No. 3.2e-44; Mismatches 189; Indels 88; Gaps 16;

Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;
 QY 62 AGTPPGR-----SPGRASPAQAS-----PARASPALASLRSSSGRSSARS 104
 Db 5 SGSPPAIGPYENHGYQENPYPAQPTVPTVEVHPAQYF----- 46
 QY 105 ASVTPSPRYLVATPVGAVPIRSSPARSAPATRATESPOTSLPKFTWEGOKQLPL- 163
 Db 47 -----SPVQYAPRVLTOASNPVCTQPKSPSGTCTSKT-----KALCIT 88
 QY 164 --IGCVLLIALVLSLILFQFWQ-----GHTGIRYKQRESCEPKAVRCDCGVDCIK 215
 Db 89 LTLGTFVLGALLAGLL-----WKFMGSKCSNSGIEC-DSSGTCINPNCWDGVSHCPGG 142
 QY 216 SDELGCVRFMDKSLIKTYSGSSHOWLPICSSNNDSYSEKTCQOLGFEASHRTTEVAHR 275
 Db 143 EDENRCVRLYGNPNTLQWYSSQKSMHVCQDDMNENYGRAACRMGKKNFYSSQ-GIV 201
 QY 276 DFANSFSLRYNST-----IOESLHRSB-CPSORYISIQCSHGRL--AMTGRIVGAL 326
 Db 202 DDSGSTFMKNTSAGNVDYKQLYHSDACSSKAVVSLRCLACGVNLSSRSQSRIVGES 261
 QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTEKEVLEG--WKVYAGTNTL 383
 Db 262 ALPGAMPQVSLHVNVAHVCSSITTPRMIVTAHCV-----EKPLNNPMHTAFAGILRQ 317
 QY 384 HOLPEAA--SIAETIINSNYDEEDDYDIALMRSLKPTLSAHNIPACLPMHGQFSLNE 441
 Db 318 SFMFYAGYQVQKISHENYDSKTKNDIALMKLOKPLTFNDLVKPVCLPMPGMLOPEQ 377
 QY 442 TCWITGFKETRETDDKTSPLREYQVNLIDPKKNDYLVYDSYLTFRMPCAGDLRGDS 501
 Db 378 LCMISGMWATEE-KGKTSEVINAALVLLIETQRCNSRYVDNLITPAMICAGFLOGNDS 436
 QY 502 CQDSSGGLVCEQNNRWYLAGVTSWGTGCGGRNKGVTYTKTEVLPWITYSKMES 555
 Db 437 CQDSSGGLVTSNNNIMWLIGDTSWGSACAKAYRPGYGVNMFWDWIYRQKA 490

RESULT 14

US-09-759-143-895
Sequence 895; Application US/09759143

Patent No. 6800746

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Devin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolck, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.

```

; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 895
; LENGTH: 492
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-759-143-895

Query Match      24.6%; Score 736.5; DB 2; Length 492;
Best Local Similarity 32.4%; Pred. No. 3.2e-44;
Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;

QY 62 AGTPGRR-----SFGSPAPQAS-----PARASPALASLSRSSSGSSSARS 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 SSGSPALGPPYENHGYOPENPYPAQPTVPTVYEVHQAQYYP----- 46
QY 105 ASVTTSPTRYLVATPVGAVPIRSSPARSAPATPATRESPTSLPKFTWREGOKPL- 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47 -----SPVQYAPRVLTAQSNPVCTOPKSPSGTCTSKT-----KKALCIT 88
QY 164 --IGCVLLIALVSLILFQFMQ-----GHTGIRYKQESCPKHAVRCDGVDCXK 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 LTLGFLVGAALAGL-----MKFMGSKCSNSGIEC-DSSGTCINPNNWCDGVSHCPG 142
QY 216 SDELGCYAFMDKSLTKYSGSSHQWLPICSSNNMDSSEKTCOOLGFEASARTTEVAHR 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 143 EDENRCVRLYGPNIILQWYSSQKSMHPCVCDNMENENGRAACRMGKNNFYSSQ-GIV 201
QY 276 DFANSFSLIRYNST-----IOESLHRSE-CPQORYISLQCSHGGR--AMTGRIYVGA 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 202 DDSGSTSMKLNKNTSAGNDIYKLYHSDACSSKAVSLRCLACGVNLNSSROSRIVGES 261
QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTREKYLEG--WKVYAGTSLN 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 262 ALPGAMPQVSLHQNVAHVCSSITTPREIVTAHCV---EKPLNPMHTATFAGILRQ 317
QY 384 HOLPEAA--SIAEIINSNTYDEEDVDIALMRLSKPTLSAHIPACLPNHGQTFSLNE 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 318 SFMEYGAQYQVQKXISHPNYDSKTKNDIALMKLOKPLTFENDLVKPVCLPMPGMLOPEQ 377
QY 442 TCWITGFGKTRTDKTSPPFLREVQVNLIDFKKCNVDYLVDSYLTTPRMWCGADLRGSDS 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 378 LCMISGMKATEE-KGKTSVULNAKVLILFQRCNSRYVDNLITPAMICAGFLOGNDS 436
QY 502 CQGDGGGLVCEQNNRWYLAGVTSWGTGCGQRRNKPGVYTKYTEVLPMYISKMS 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 CQGDGGGLVTSNNNIMWLIGDTSMGSGCAYRPGVYGNVFTDWIYRQKA 490

RESULT 15
US-10-012-896-895
; Sequence 895, Application US/10012896
; Patent No. 6943236
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
```

```

; APPLICANT: Vedick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassole, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 895
; LENGTH: 492
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-012-896-895

Query Match      24.6%; Score 736.5; DB 2; Length 492;
Best Local Similarity 32.4%; Pred. No. 3.2e-44;
Matches 173; Conservative 84; Mismatches 189; Indels 88; Gaps 16;

QY 62 AGTPGRR-----SFGSPAPQAS-----PARASPALASLSRSSSGSSSARS 104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5 SSGSPALGPPYENHGYOPENPYPAQPTVPTVYEVHQAQYYP----- 46
QY 105 ASVTTSPTRYLVATPVGAVPIRSSPARSAPATPATRESPTSLPKFTWREGOKPL- 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47 -----SPVQYAPRVLTAQSNPVCTOPKSPSGTCTSKT-----KKALCIT 88
QY 164 --IGCVLLIALVSLILFQFMQ-----GHTGIRYKQESCPKHAVRCDGVDCXK 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 LTLGFLVGAALAGL-----MKFMGSKCSNSGIEC-DSSGTCINPNNWCDGVSHCPG 142
QY 216 SDELGCYAFMDKSLTKYSGSSHQWLPICSSNNMDSSEKTCOOLGFEASARTTEVAHR 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 143 EDENRCVRLYGPNIILQWYSSQKSMHPCVCDNMENENGRAACRMGKNNFYSSQ-GIV 201
QY 276 DFANSFSLIRYNST-----IOESLHRSE-CPQORYISLQCSHGGR--AMTGRIYVGA 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 202 DDSGSTSMKLNKNTSAGNDIYKLYHSDACSSKAVSLRCLACGVNLNSSROSRIVGES 261
QY 327 ASDSKMPQVSLHFGTTHICGTLIDAQWVLTAAHCFVTREKYLEG--WKVYAGTSLN 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 262 ALPGAMPQVSLHQNVAHVCSSITTPREIVTAHCV---EKPLNPMHTATFAGILRQ 317
QY 384 HOLPEAA--SIAEIINSNTYDEEDVDIALMRLSKPTLSAHIPACLPNHGQTFSLNE 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 318 SFMEYGAQYQVQKXISHPNYDSKTKNDIALMKLOKPLTFENDLVKPVCLPMPGMLOPEQ 377
QY 442 TCWITGFGKTRTDKTSPPFLREVQVNLIDFKKCNVDYLVDSYLTTPRMWCGADLRGSDS 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 378 LCMISGMKATEE-KGKTSVULNAKVLILFQRCNSRYVDNLITPAMICAGFLOGNDS 436
QY 502 CQGDGGGLVCEQNNRWYLAGVTSWGTGCGQRRNKPGVYTKYTEVLPMYISKMS 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 CQGDGGGLVTSNNNIMWLIGDTSMGSGCAYRPGVYGNVFTDWIYRQKA 490

Search completed: September 16, 2006, 03:15:05
Job time : 52 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:25:36 ; Search time 181 Seconds
(without alignments)
1438.269 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999

Sequence: 1 MERSHGNASPARTPSAGAS.....TEVLPMTYSKMSERFRKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/us11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	100.0	562	3	US-09-879-792-12 Sequence 12, Appl
2	2999	100.0	562	4	US-10-806-370-12 Sequence 12, Appl
3	2999	100.0	569	4	US-10-428-275-130 Sequence 130, App
4	2991	99.7	562	4	US-10-156-214A-24 Sequence 24, Appl
5	2991	99.7	562	5	US-10-916-758-94 Sequence 94, Appl
6	2963	98.8	688	4	US-10-156-214A-26 Sequence 26, Appl
7	2963	98.8	688	5	US-10-916-758-2 Sequence 2, Appl
8	2956	98.6	581	4	US-10-353-690-100 Sequence 100, App
9	2948.5	98.3	586	4	US-10-428-275-104 Sequence 104, App
10	2948.5	98.3	586	4	US-10-428-275-124 Sequence 124, App
11	2948.5	98.3	593	4	US-10-428-275-120 Sequence 120, App
12	2943.5	98.1	586	4	US-10-428-275-142 Sequence 142, App
13	2942.5	98.1	586	4	US-10-428-275-140 Sequence 140, App
14	2775.5	92.5	537	3	US-09-888-615-104 Sequence 104, App
15	2775.5	92.5	537	4	US-10-428-275-128 Sequence 128, App
16	2775.5	92.5	537	6	US-11-037-243-104 Sequence 104, App
17	2686	89.6	542	4	US-10-428-275-122 Sequence 122, App
18	2291.5	76.4	477	4	US-10-177-661-2 Sequence 2, Appl
19	2291.5	76.4	477	5	US-09-898-837A-15 Sequence 2, Appl
20	2287	76.3	421	3	US-09-898-837A-15 Sequence 15, Appl
21	2287	76.3	421	4	US-10-428-275-132 Sequence 132, App
22	2219.5	74.0	486	4	US-10-274-639-8 Sequence 8, Appl
23	2219.5	74.0	486	4	US-10-333-574-8 Sequence 8, Appl
24	2130.5	71.0	412	4	US-10-428-275-144 Sequence 144, App
25	2108	70.3	446	4	US-10-177-661-4 Sequence 4, Appl
26	2108	70.3	446	5	US-10-910-507-4 Sequence 4, Appl
27	2097	69.9	382	4	US-10-428-275-134 Sequence 134, App

28	2059	68.7	406	4	US-10-428-275-126 Sequence 126, App
29	2059	68.7	406	4	US-10-428-275-138 Sequence 138, App
30	2026	67.6	401	4	US-10-428-275-112 Sequence 112, App
31	2016	67.2	401	4	US-10-428-275-116 Sequence 116, App
32	2012	67.1	401	4	US-10-428-275-118 Sequence 118, App
33	1842.5	61.4	370	4	US-10-428-275-114 Sequence 114, App
34	1839.5	61.3	367	4	US-10-428-275-110 Sequence 110, App
35	1311	43.7	296	3	US-09-804-156-10 Sequence 30, Appl
36	1311	43.7	296	4	US-10-067-761-30 Sequence 30, Appl
37	1311	43.7	296	4	US-10-319-519-30 Sequence 30, Appl
38	1311	43.7	372	3	US-09-804-156-15 Sequence 15, Appl
39	1311	43.7	372	3	US-09-946-633-7 Sequence 7, Appl
40	1311	43.7	372	4	US-10-125-459-7 Sequence 7, Appl
41	1311	43.7	372	4	US-10-067-761-15 Sequence 15, Appl
42	1311	43.7	372	4	US-10-319-519-15 Sequence 15, Appl
43	1270	42.3	239	4	US-10-428-275-106 Sequence 106, App
44	1270	42.3	239	4	US-10-428-275-108 Sequence 108, App
45	1265	42.2	231	4	US-10-428-275-136 Sequence 136, App

ALIGNMENTS

```
RESULT 1
US-09-879-792-12
; Sequence 12, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE REFERENCE: 02973.00035
; CURRENT FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-879-792-12
Query Match 100.0%; Score 2999; DB 3; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MERSHGNASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
1 MERSHGNASPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
61 PAGTPPGRASGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
61 PAGTPPGRASGRASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 120
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWRBQKQPLIGCVLLIALVSLIL 180
121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWRBQKQPLIGCVLLIALVSLIL 180
181 FQFQGHGTGIRYKQRESCEPHAVRCQGVNDCKLKSDELGVREDMKSLIKIYSGSSHQ 240
181 FQFQGHGTGIRYKQRESCEPHAVRCQGVNDCKLKSDELGVREDMKSLIKIYSGSSHQ 240
241 WLPICSSNMNDVSEKTCQQLGFESARHTEVAHRDFANSFSILRYNSTIOESLHRSCEP 300
```

Db 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Qy 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAOVLTAA 360
Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAOVLTAA 360
Qy 361 HCFVTRKVLKGVKVAAGTSLHQLPEAASIAEIIINSNTDEEDDYDIALMLSKPLT 420
Db 361 HCFVTRKVLKGVKVAAGTSLHQLPEAASIAEIIINSNTDEEDDYDIALMLSKPLT 420
Qy 421 LSAIHHPACLPMHQOTFSLNCTWITGFKTRTDDKTSPLREVOVNLIDFKKCNVLY 480
Db 421 LSAIHHPACLPMHQOTFSLNCTWITGFKTRTDDKTSPLREVOVNLIDFKKCNVLY 480
Qy 481 YDSYLTRMMACGLRGGRDSCQDSCGGLVCEONNRWYLAGVTSWGTGCGQRKPGVYT 540
Db 481 YDSYLTRMMACGLRGGRDSCQDSCGGLVCEONNRWYLAGVTSWGTGCGQRKPGVYT 540
Qy 541 KTEVLPWISKMESEVFRKS 562
Db 541 KTEVLPWISKMESEVFRKS 562

RESULT 2

US-10-806-370-12
Sequence 12, Application US/10806370
Publication No. US20040209327A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonghong
TITLE OF INVENTION: Regulation of Human Transmembrane Serine
FILE REFERENCE: 02973.00035
CURRENT APPLICATION NUMBER: US/10/806.370
CURRENT FILING DATE: 2004-03-23
PRIOR APPLICATION NUMBER: US/09/879.792
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/211,224
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 60/283,353
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/283,648
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: PCT _____ (Docket No. LIO-81-WO)
PRIOR FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-10-806-370-12

Query Match 100.0%; Score 2999; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 1 MERDSHGNSAPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
Qy 61 PAGTPPGRASPGASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Db 61 PAGTPPGRASPGASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Qy 121 PVGAVPIRSSPARAPATRATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180
Db 121 PVGAVPIRSSPARAPATRATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180
Qy 181 FQFMQHTGIIYKXORSCPKHAYRCGVVDCXKSDGLGCVRFDMDSLLIKTISGSSHQ 240
Db 181 FQFMQHTGIIYKXORSCPKHAYRCGVVDCXKSDGLGCVRFDMDSLLIKTISGSSHQ 240

Qy 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Db 241 WLPICSSNMNDYSSEKTCQOLGFESARHTEVAHRDFANSFSLRYNSTIOESLHRSBCP 300
Qy 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAOVLTAA 360
Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAOVLTAA 360
Qy 361 HCFVTRKVLKGVKVAAGTSLHQLPEAASIAEIIINSNTDEEDDYDIALMLSKPLT 420
Db 361 HCFVTRKVLKGVKVAAGTSLHQLPEAASIAEIIINSNTDEEDDYDIALMLSKPLT 420
Qy 421 LSAIHHPACLPMHQOTFSLNCTWITGFKTRTDDKTSPLREVOVNLIDFKKCNVLY 480
Db 421 LSAIHHPACLPMHQOTFSLNCTWITGFKTRTDDKTSPLREVOVNLIDFKKCNVLY 480
Qy 481 YDSYLTRMMACGLRGGRDSCQDSCGGLVCEONNRWYLAGVTSWGTGCGQRKPGVYT 540
Db 481 YDSYLTRMMACGLRGGRDSCQDSCGGLVCEONNRWYLAGVTSWGTGCGQRKPGVYT 540
Qy 541 KTEVLPWISKMESEVFRKS 562
Db 541 KTEVLPWISKMESEVFRKS 562

RESULT 3

US-10-428-275-130
Sequence 130, Application US/10428275
Publication No. US20040067505A1
GENERAL INFORMATION:
APPLICANT: Alvarez et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-585
CURRENT APPLICATION NUMBER: US/10/428.275
CURRENT FILING DATE: 2003-05-01
PRIOR APPLICATION NUMBER: 09/966545
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/544511
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128514
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: 09/569269
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/134315
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/619252
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/789390
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/185548
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 450
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 130
LENGTH: 569
TYPE: PRT
ORGANISM: Homo sapiens
US-10-428-275-130

Query Match 100.0%; Score 2999; DB 4; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERDSHGNSAPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 5 MERDSHGNSAPARTPSAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 64
Qy 61 PAGTPPGRASPGASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 120
Db 65 PAGTPPGRASPGASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLVAT 124
Qy 121 PVGAVPIRSSPARAPATRATRESPTSLPKFTWREGKQPLIGCVLLILALVVSLLITL 180

Db 125 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVVSLLIL 184
Qy 181 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 240
Db 185 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 244
Qy 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 245 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 304
Qy 301 SQRYSIQCSHCGIRAMTGRIVGALASDSKMPQVSLHFCTTHICGCTLLDAQWVLTAA 360
Db 305 SQRYSIQCSHCGIRAMTGRIVGALASDSKMPQVSLHFCTTHICGCTLLDAQWVLTAA 364
Qy 361 HCFEVTREKYLEGKVVAGTNSLHQLPEAASIAEIIINSNTDEDDYDIALMRLSKPLT 420
Db 365 HCFEVTREKYLEGKVVAGTNSLHQLPEAASIAEIIINSNTDEDDYDIALMRLSKPLT 424
Qy 421 LSAHIHPACLPMHGQTSINETCMTGFGKTRTDDKTSPLREVOVNLIDFKKCNBYLV 480
Db 425 LSAHIHPACLPMHGQTSINETCMTGFGKTRTDDKTSPLREVOVNLIDFKKCNBYLV 484
Qy 481 YDSYLTFRMCAADLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRKPKGYTT 540
Db 485 YDSYLTFRMCAADLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRKPKGYTT 544
Qy 541 KTEVLPMTIYSKMESEVFRKS 562
Db 545 KTEVLPMTIYSKMESEVFRKS 566

RESULT 4

US-10-156-214A-24
; Sequence 24, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempile
; APPLICANT: George P. Vlasuk
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Siev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 562
; TYPE: PRT
; ORGANISM: homo sapien
; FEATURE:
; NAME/KEY: protease domain of endochelase 2
; LOCATION: (321)..(562)
US-10-156-214A-24

Query Match 99.7%; Score 2991; DB 4; Length 562;
Best Local Similarity 99.8%; Pred. No. 7,2e-178;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MERDHSNAGSPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 1 MERDHSNAGSPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Qy 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Db 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Qy 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVVSLLIL 180
Db 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVVSLLIL 180

Qy 181 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 240
Db 181 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 240
Qy 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDVSEKTCQQLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Qy 301 SQRYSIQCSHCGIRAMTGRIVGALASDSKMPQVSLHFCTTHICGCTLLDAQWVLTAA 360
Db 301 SQRYSIQCSHCGIRAMTGRIVGALASDSKMPQVSLHFCTTHICGCTLLDAQWVLTAA 360
Qy 361 HCFEVTREKYLEGKVVAGTNSLHQLPEAASIAEIIINSNTDEDDYDIALMRLSKPLT 420
Db 361 HCFEVTREKYLEGKVVAGTNSLHQLPEAASIAEIIINSNTDEDDYDIALMRLSKPLT 420
Qy 421 LSAHIHPACLPMHGQTSINETCMTGFGKTRTDDKTSPLREVOVNLIDFKKCNBYLV 480
Db 421 LSAHIHPACLPMHGQTSINETCMTGFGKTRTDDKTSPLREVOVNLIDFKKCNBYLV 480
Qy 481 YDSYLTFRMCAADLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRKPKGYTT 540
Db 481 YDSYLTFRMCAADLRGGRDSCQDSSGGLVCEQNNRWYLAGVTSWGTGCGQRKPKGYTT 540
Qy 541 KTEVLPMTIYSKMESEVFRKS 562
Db 541 KTEVLPMTIYSKMESEVFRKS 562

RESULT 5

US-10-916-758-94
; Sequence 94, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-94

Query Match 99.7%; Score 2991; DB 5; Length 562;
Best Local Similarity 99.8%; Pred. No. 7,2e-178;
Matches 561; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MERDHSNAGSPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Db 1 MERDHSNAGSPARTPSAGASPAQASPACTPPGRASPAQASPAQASPAQASPAQAS 60
Qy 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Db 61 PACTPPGRASPGRASPAQASPARASPALASLSRSSSGSSARSASVTTSTRTVYLVAT 120
Qy 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVVSLLIL 180
Db 121 PVGAVPIRSSPARSAPATRAATRESPTSLPKFTWREGQKQLPLIGCVLLIALVVSLLIL 180
Qy 181 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 240
Db 181 FQWQGHGTGIRYKQORSSCPKHAIVRCGVVDCDKLSDELGCVREDDMDKSLIKITYSGSSHQ 240

QY 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
QY 301 SORVYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOVLTAA 360
Db 301 SORVYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOVLTAA 360
QY 361 HCFVYTRREKVLGKMWKVVAGTSMHLQLPAAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
Db 361 HCFVYTRREKVLGKMWKVVAGTSMHLQLPAAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
QY 421 LSAHIHPACLPMHQGTFSINETCWTGFGKTRTDDKTSPLREYQVNLIDFKCNDLV 480
Db 421 LSAHIHPACLPMHQGTFSINETCWTGFGKTRTDDKTSPLREYQVNLIDFKCNDLV 480
QY 481 YDSYLTTPMMCAGLRGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
Db 481 YDSYLTTPMMCAGLRGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
QY 541 KTEVLPMIYSKMESEVRFRKS 562
Db 541 KTEVLPMIYSKMESEVRFRKS 562

RESULT 6

US-10-156-214A-26
; Sequence 26, Application US/10156214A
; Publication No. US20040001801A1
; GENERAL INFORMATION:
; APPLICANT: Edwin L. Madison
; APPLICANT: Joseph Edward Sempile
; APPLICANT: George P. Vlahek
; APPLICANT: Scott Jeffrey Kemp
; APPLICANT: Mallareddy Komandla
; APPLICANT: Daniel Vanna Slev
; TITLE OF INVENTION: Conjugates Activated By Cell Surface Proteases and Therapeutic Use
; TITLE OF INVENTION: Therapeutic Use
; FILE REFERENCE: 24745-1611
; CURRENT APPLICATION NUMBER: US/10/156,214A
; CURRENT FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: protease domain
; LOCATION: (321)..(688)
US-10-156-214A-26

Query Match 98.8%; Score 2963; DB 4; Length 688;
Best Local Similarity 99.3%; Pred. No. 4.9e-176;

Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNASPARTPSAGASPAQASPAQTTPGRASPAQASPAQASPAQTTPGRASPAQAS 60
Db 1 MERDSHGNASPARTPSAGASPAQASPAQTTPGRASPAQASPAQASPAQTTPGRASPAQAS 60
QY 61 PAGTTPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLYVRAT 120
Db 61 PAGTTPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLYVRAT 120
QY 121 PVGAVPIRSSPARAPATRAATRESPTGSLPKFTWRBQOKPLIGCVLLIALVSLITL 180
Db 121 PVGAVPIRSSPARAPATRAATRESPTGSLPKFTWRBQOKPLIGCVLLIALVSLITL 180
QY 181 FQFMOGHTGIRYKQRESCKHAVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQFMOGHTGIRYKQRESCKHAVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
QY 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300

Db 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
QY 301 SORVYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOVLTAA 360
Db 301 SORVYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOVLTAA 360
QY 361 HCFVYTRREKVLGKMWKVVAGTSMHLQLPAAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
Db 361 HCFVYTRREKVLGKMWKVVAGTSMHLQLPAAASIAEIIINSNTYDEEDDYDIALMRLSKPLT 420
QY 421 LSAHIHPACLPMHQGTFSINETCWTGFGKTRTDDKTSPLREYQVNLIDFKCNDLV 480
Db 421 LSAHIHPACLPMHQGTFSINETCWTGFGKTRTDDKTSPLREYQVNLIDFKCNDLV 480
QY 481 YDSYLTTPMMCAGLRGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
Db 481 YDSYLTTPMMCAGLRGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNKPgyt 540
QY 541 KTEVLPMIYSKMESEVR 558
Db 541 KTEVLPMIYSKMESEVR 558

RESULT 7

US-10-916-758-2
; Sequence 2, Application US/10916758
; Publication No. US20050180977A1
; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Madison, Edwin L.
; TITLE OF INVENTION: ENDOTHELIAL-2 LIGANDS
; FILE REFERENCE: 10280-065001
; CURRENT APPLICATION NUMBER: US/10/916,758
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/520,164
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/495,005
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-916-758-2

Query Match 98.8%; Score 2963; DB 5; Length 688;
Best Local Similarity 99.3%; Pred. No. 4.9e-176;
Matches 554; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERDSHGNASPARTPSAGASPAQASPAQTTPGRASPAQASPAQASPAQTTPGRASPAQAS 60
Db 1 MERDSHGNASPARTPSAGASPAQASPAQTTPGRASPAQASPAQASPAQTTPGRASPAQAS 60
QY 61 PAGTTPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLYVRAT 120
Db 61 PAGTTPGRASPGRASPAQASPARASPALASLSRSSSGRSSSARASAVTTSPTRYLYVRAT 120
QY 121 PVGAVPIRSSPARAPATRAATRESPTGSLPKFTWRBQOKPLIGCVLLIALVSLITL 180
Db 121 PVGAVPIRSSPARAPATRAATRESPTGSLPKFTWRBQOKPLIGCVLLIALVSLITL 180
QY 181 FQFMOGHTGIRYKQRESCKHAVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
Db 181 FQFMOGHTGIRYKQRESCKHAVRCGVVDCKLSDELGCVRFDMDKSLIKIYSGSSHQ 240
QY 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMNDYSSEKTCQOOLGFESAHRTTEVAHRDPANSPSILRYNSTIOESLHRSCEP 300
QY 301 SORVYISLQCSHCGRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAOVLTAA 360

Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQWVLTAA 360
QY HCFVTRKVLGKMKVYAGTSLNHLQPEASIAEIIINSNTDEDDVDIALMRSLKPLT 420
Db 361 HCFVTRKVLGKMKVYAGTSLNHLQPEASIAEIIINSNTDEDDVDIALMRSLKPLT 420
QY 421 LSAIHHPACLPFHGOTFSNETCWTITGKTRTDDKTSPLREVQVNLIDFKKNDLV 480
Db 421 LSAIHHPACLPFHGOTFSNETCWTITGKTRTDDKTSPLREVQVNLIDFKKNDLV 480
QY 481 YDSVLTTPMAGDLRGGRDSCQDGSGLVCEQNNRWYLAGVTSWGTCCGQRNKPYYT 540
Db 481 YDSVLTTPMAGDLRGGRDSCQDGSGLVCEQNNRWYLAGVTSWGTCCGQRNKPYYT 540
QY 541 KTEVLPWYSKMESEVR 558
Db 541 KTEVLPWYSKMESEVR 558
RESULT 8
US-10-353-690-100
; Sequence 100, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 84624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 2686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; FILE REFERENCE: MP102-018P1RMONIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/386,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-353-690-100

Query Match 98.6%; Score 2956; DB 4; Length 581;
Best Local Similarity 99.8%; Pred. No. 1,1e-175;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
Db 1 MERDSHGASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
QY 61 PACTPPGRASPGRASPAQASPARASPALASISRSSSSGSSSARASVTSSTRVYLVAT 120
Db 61 PACTPPGRASPGRASPAQASPARASPALASISRSSSSGSSSARASVTSSTRVYLVAT 120
QY 121 PVGAVPIRSSPARASAPTRATRESPTSLPKFTREGOKPLIGCVULLALVYSLTL 180
Db 121 PVGAVPIRSSPARASAPTRATRESPTSLPKFTREGOKPLIGCVULLALVYSLTL 180
QY 181 PQFQGHGTGIRYKQRESCPHVAVRCDGVYDCKLSDBLGCVRPDMKSLIKIYSSSHQ 240
Db 181 PQFQGHGTGIRYKQRESCPHVAVRCDGVYDCKLSDBLGCVRPDMKSLIKIYSSSHQ 240
QY 241 WLPICSSNMWDSYSEKTCQQLGFSSAHTTVAARDPANSFSLIRYNSTIOESLHRSCEP 300
Db 241 WLPICSSNMWDSYSEKTCQQLGFSSAHTTVAARDPANSFSLIRYNSTIOESLHRSCEP 300
QY 301 SQRYSIQCSHCGRAMTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQWVLTAA 360
Db 301 SQRYSIQCSHCGRAMTGRIVGALASDSKPMQVSLHFGTTHICGTLIDAQWVLTAA 360
QY 361 HCFVTRKVLGKMKVYAGTSLNHLQPEASIAEIIINSNTDEDDVDIALMRSLKPLT 420
Db 361 HCFVTRKVLGKMKVYAGTSLNHLQPEASIAEIIINSNTDEDDVDIALMRSLKPLT 420
QY 421 LSAIHHPACLPFHGOTFSNETCWTITGKTRTDDKTSPLREVQVNLIDFKKNDLV 480
Db 421 LSAIHHPACLPFHGOTFSNETCWTITGKTRTDDKTSPLREVQVNLIDFKKNDLV 480
QY 481 YDSVLTTPMAGDLRGGRDSCQDGSGLVCEQNNRWYLAGVTSWGTCCGQRNKPYYT 540
Db 481 YDSVLTTPMAGDLRGGRDSCQDGSGLVCEQNNRWYLAGVTSWGTCCGQRNKPYYT 540
QY 541 KTEVLPWYSKME 554
Db 541 KTEVLPWYSKME 554
RESULT 9
US-10-428-275-104
; Sequence 104, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428,275
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450

SOFTWARE: Curaseqblast version 0.1
 SEQ ID NO 104
 LENGTH: 586
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-428-275-104

Query Match 98.3%; Score 2948.5; DB 4; Length 586;
 Best Local Similarity 99.1%; Pred. No. 3.3e-175;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

QY 1 MERDSHGNAAPKRTSPASAPAOASPAGTPPGRASPAOASPAGTPPGRASPAOAS 60
DB 1 MERDSHGNAAPKRTSPASAPAOASPAGTPPGRASPAOASPAGTPPGRASPAOAS 60
QY 61 PAGTPGRASPGR-----ASPAOASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 115
DB 61 PAGTPGRASPGRASPAOASPAPASPAOASPAOASPAOASPAOASPAOASPAOAS 120
QY 116 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 175
DB 121 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 180
QY 176 SLIIIFQWOGHTGIRYKQRESCKPAVRCGVVDCDKLKSDELGCVRFDMDKSLIKIYS 235
DB 181 SLIIIFQWOGHTGIRYKQRESCKPAVRCGVVDCDKLKSDELGCVRFDMDKSLIKIYS 240
QY 236 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 295
DB 241 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 300
QY 296 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAQW 355
DB 301 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAQW 360
QY 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLPEASIAEIIINSYTTBEDDYDIALMRL 415
DB 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLPEASIAEIIINSYTTBEDDYDIALMRL 420
QY 416 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDPFKC 475
DB 421 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDPFKC 480
QY 476 NDLYVDSYLTFRMWCAGDLRGGRSCGDSGGPLVCEQNNRWYLAGVTSMTGCGGRNK 535
DB 481 NDLYVDSYLTFRMWCAGDLRGGRSCGDSGGPLVCEQNNRWYLAGVTSMTGCGGRNK 540
QY 536 PGVYTKVTEVLPWYISKME 554
DB 541 PGVYTKVTEVLPWYISKME 559

```

RESULT 10
 US-10-428-275-124
 ; Sequence 124, Application US/10428275
 ; Publication No. US20040067505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alvarez et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 ; FILE REFERENCE: 21402-585
 ; CURRENT APPLICATION NUMBER: US/10/428,275
 ; PRIOR FILING DATE: 2003-05-01
 ; PRIOR APPLICATION NUMBER: 09/966545
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 09/544511
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: 60/128514
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 09/569269
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: 60/134315
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: 09/619252

PRIOR FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/789390
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/185548
 PRIOR FILING DATE: 2000-02-25
 NUMBER OF SEQ ID NOS: 450
 SOFTWARE: Curaseqblast version 0.1
 SEQ ID NO 124
 LENGTH: 586
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-428-275-124

Query Match 98.3%; Score 2948.5; DB 4; Length 586;
 Best Local Similarity 99.1%; Pred. No. 3.3e-175;
 Matches 554; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

QY 1 MERDSHGNAAPKRTSPASAPAOASPAGTPPGRASPAOASPAGTPPGRASPAOAS 60
DB 1 MERDSHGNAAPKRTSPASAPAOASPAGTPPGRASPAOASPAGTPPGRASPAOAS 60
QY 61 PAGTPGRASPGR-----ASPAOASPARASPALASLSRSSSGRSSSARSASAVTTSPTRY 115
DB 61 PAGTPGRASPGRASPAOASPAPASPAOASPAOASPAOASPAOASPAOASPAOAS 120
QY 116 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 175
DB 121 LVRATPVGAVPIRSSPARASAPATRAATRESPTSLPKFTWREGQQLPLIGCVLLIALV 180
QY 176 SLIIIFQWOGHTGIRYKQRESCKPAVRCGVVDCDKLKSDELGCVRFDMDKSLIKIYS 235
DB 181 SLIIIFQWOGHTGIRYKQRESCKPAVRCGVVDCDKLKSDELGCVRFDMDKSLIKIYS 240
QY 236 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 295
DB 241 GSSHOMLPICSNMNDYSSEKTCQQLGFESAHRTTEVAHRDPANFSILRYNSTIOESLH 300
QY 296 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAQW 355
DB 301 RSECPQRYISLQCSHCGLRAMTGRIVGALASDSKWPQVSLHFGTTHICGTLIDAQW 360
QY 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLPEASIAEIIINSYTTBEDDYDIALMRL 415
DB 361 VLTAAHCFVTRREKYLEGMKYAGTSLNHLPEASIAEIIINSYTTBEDDYDIALMRL 420
QY 416 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDPFKC 475
DB 421 SKPLTSLAHIPACLPMHGOTFSLNETCMITGFGKTRTDKTSPLREVOVNLIDPFKC 480
QY 476 NDLYVDSYLTFRMWCAGDLRGGRSCGDSGGPLVCEQNNRWYLAGVTSMTGCGGRNK 535
DB 481 NDLYVDSYLTFRMWCAGDLRGGRSCGDSGGPLVCEQNNRWYLAGVTSMTGCGGRNK 540
QY 536 PGVYTKVTEVLPWYISKME 554
DB 541 PGVYTKVTEVLPWYISKME 559

```

RESULT 11
 US-10-428-275-120
 ; Sequence 120, Application US/10428275
 ; Publication No. US20040067505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alvarez et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
 ; FILE REFERENCE: 21402-585
 ; CURRENT APPLICATION NUMBER: US/10/428,275
 ; PRIOR FILING DATE: 2003-05-01
 ; PRIOR APPLICATION NUMBER: 09/966545
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 09/544511
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: 60/128514

Query Match	Similarity	98.1%	Score	2943.5	DB 4	Length	566
Best Local	Similarity	98.9%	Pred. No.	6.8e-175			
Matches	553	Conservative	0	Mismatches	1	Indels	5
							Gaps
							1
QY	1	MERDSHGNASDPARPPASGASPAQAS	PAGT	PPGRASPAQASPAQAS	PACT	PPGRASPAQAS	60
DB	1	MERSHGNASPARPPASGASPAQAS	PACT	PPGRASPAQASPAQAS	PACT	PPGRASPAQAS	60
QY	61	PAGT	PPGRASFCR	-----ASPAQAS	SPARAS	PALLASLS	115
DB	61	PAGT	PPGRASFCR	SPARAS	PAQASPAQAS	PALASLS	120
QY	116	LVRATPVGAIVIRSSPARAS	PATRAT	TRSPGTS	PKPTWR	REGOKPL	175
DB	121	LVRATPVGAIVIRSSPARAS	PATRAT	TRSPGTS	PKPTWR	REGOKPL	180
QY	176	SLIL	FOFMOGHTGIRYKEORE	SECPKAVR	CDGVVDC	LKSDELGCVR	235
DB	181	SLIL	FOFMOGHTGIRYKEORE	SECPKAVR	CDGVVDC	LKSDELGCVR	240
QY	236	GSSHOMLPICSSNNNDYS	SEKTCQOL	GPESAHRTTEVA	HRDPANS	FSLIRNST	295
DB	241	GSSHOMLPICSSNNNDYS	SEKTCQOL	GPESAHRTTEVA	HRDPANS	FSLIRNST	300
QY	296	RSECSORYISL	OCSHGCLRAMTGI	IVGALAS	DSKWPQVSL	HLHGTHI	355
DB	301	RSECSORYISL	OCSHGCLRAMTGI	IVGALAS	DSKWPQVSL	HLHGTHI	360
QY	356	VLTAAHCFVTR	REKYLEGKMVYAGT	SNLHQLPEAAS	IAEIIINS	NYTDEBDY	415
DB	361	VLTAAHCFVTR	REKYLEGKMVYAGT	SNLHQLPEAAS	IAEIIINS	NYTDEBDY	420
QY	416	SKPLTSLAHIPAC	LPMTGOTFSLNET	WTWITGFG	TRTRTD	DTSPFL	475
DB	421	SKPLTSLAHIPAC	LPMTGOTFSLNET	WTWITGFG	TRTRTD	DTSPFL	480
QY	476	NDLYVYDSYL	LPMMACAGDL	RGRDSCGDS	GSGPLV	CEONNRWY	535
DB	481	NDLYVYDSYL	LPMMACAGDL	RGRDSCGDS	GSGPLV	CEONNRWY	540
QY	536	PGVYTKVTEVL	PMIYSKME	554			
DB	541	PGVYTKVTEVL	PMIYSKME	559			

```
/ Publication No. US20040067505A1
/ GENERAL INFORMATION:
/ APPLICANT: Alvarez et al.
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
/ FILE REFERENCE: 21402-585
/ CURRENT FILING DATE: 2003-05-01
/ PRIOR APPLICATION NUMBER: US/10/428,275
/ PRIOR FILING DATE: 2003-05-01
/ PRIOR APPLICATION NUMBER: 09/966545
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 09/544511
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: 60/128514
/ PRIOR FILING DATE: 1999-04-09
/ PRIOR APPLICATION NUMBER: 09/569269
/ PRIOR FILING DATE: 2000-05-11
/ PRIOR APPLICATION NUMBER: 60/134315
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/619252
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/789390
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/185548
/ PRIOR FILING DATE: 2000-02-25
/ NUMBER OF SEQ ID NOS: 450
/ SOFTWARE: CuraseqList version 0.1
/ SEQ ID NO 140
/ LENGTH: 586
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-428-275-140
```

```
Query Match      98.1% Score 2942.5; DB 4; Length 586;
Best Local Similarity 98.9% Pred. 7.8e-175;
Matches 553; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 MERSHGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
DB 1 MERSHGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
QY 61 PACTPPGRASPCR-----ASPAQASPARASPALASLSRSSSGRSSASASATTSPTVY 115
DB 61 PACTPPGRASPCRASPAQASPAQASPARASPALASLSRSSSGRSSASASATTSPTVY 120
QY 116 LVRAIPVAVPIRSSPARASAPATRAATRESPTSLPKFTMRBQKQLPLIGCVLLIALV 175
DB 121 LVRAIPVAVPIRSSPARASAPATRAATRESPTSLPKFTMRBQKQLPLIGCVLLIALV 180
QY 176 SLILIFQWQHTGIRYKEQRESCKHAIVRCGVVDCUKLSDDELGCVRFDMDKSLIKYS 235
DB 181 SLILIFQWQHTGIRYKEQRESCKHAIVRCGVVDCUKLSDDELGCVRFDMDKSLIKYS 240
QY 236 GSSHOMLPICSSNMWDSYSEKTCQOLGFESAHRTEVVAHRDPFANSFILRYNSTIQESLH 295
DB 241 GSSHOMLPICSSNMWDSYSEKTCQOLGFESAHRTEVVAHRDPFANSFILRYNSTIQESLH 300
QY 296 RSECPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAQW 355
DB 301 RSECPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGLIDAQW 360
QY 356 VLTAAHCFVTRREKLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDIALMRL 415
DB 361 VLTAAHCFVTRREKLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDIALMRL 420
QY 416 SKPLTSLAHIPACLPMHGQTFSLNETCMTITGFKTRTDDTSPFLREVQVNLIDFKKC 475
DB 421 SKPLTSLAHIPACLPMHGQTFSLNETCMTITGFKTRTDDTSPFLREVQVNLIDFKKC 480
QY 476 NDVLYVDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNK 535
DB 481 NDVLYVDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGCGQRNK 540
QY 536 PGVYTKVTEVLPWYISKME 554
DB 540 PGVYTKVTEVLPWYISKME 554
```

```
DB 541 PGVYTKVTEVLPWYISKME 559
RESULT 14
US-09-888-615-104
/ Sequence 104, Application US/09888615
/ Patent No. US20020064856A1
/ GENERAL INFORMATION:
/ APPLICANT: PLOWMAN, GREGORY
/ APPLICANT: MAYHE, DAVID
/ APPLICANT: CAENEPEEL, SEAN
/ APPLICANT: MANNING, GERARD
/ APPLICANT: SUDARSANAM, SUCHA
/ TITLE OF INVENTION: NOVEL PROTEASES
/ FILE REFERENCE: 038602/1214
/ CURRENT APPLICATION NUMBER: US/09/888,615
/ PRIOR FILING DATE: 2001-06-26
/ PRIOR APPLICATION NUMBER: 60/214,047
/ PRIOR FILING DATE: 2000-06-26
/ NUMBER OF SEQ ID NOS: 150
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 104
/ LENGTH: 537
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-888-615-104
```

```
Query Match      92.5% Score 2775.5; DB 3; Length 537;
Best Local Similarity 92.0% Pred. No. 1.8e-164;
Matches 526; Conservative 0; Mismatches 1; Indels 45; Gaps 2;

QY 1 MERSHGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
DB 1 MERSHGNASPARTPSAGASPAQASPAQTPPGRASPAQASPAQASPAQTPPGRASPAQAS 60
QY 61 PACTPPGRASPCR-----ASPAQASPARASPALASLSRSSSGRSSASARASATTS 110
DB 61 PACTPPGRASPCRASPAQASPAQASPARASPALASLSRSSSGRSSASARASATTS 120
QY 111 PTRYIVRAIPVAVPIRSSPARASAPATRAATRESPTSLPKFTMRBQKQLPLIGCVLL 170
DB 121 PTRYIVRAIPVAVPIRSSPARASAPATRAATRESPTSLPKFTMRBQKQLPLIGCVLL 155
QY 171 IALVSLIILFQWQHTGIRYKEQRESCKHAIVRCGVVDCUKLSDDELGCVRFDMDKSL 230
DB 156 IALVSLIILFQWQHTGIRYKEQRESCKHAIVRCGVVDCUKLSDDELGCVRFDMDKSL 205
QY 231 LKIYSGSSHOMLPICSSNMWDSYSEKTCQOLGFESAHRTEVVAHRDPFANSFILRYNSTI 290
DB 206 LKIYSGSSHOMLPICSSNMWDSYSEKTCQOLGFESAHRTEVVAHRDPFANSFILRYNSTI 265
QY 291 QESLHRSRCPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGL 350
DB 266 QESLHRSRCPSORIYISLQCSHCGIRAMTGRIVGALASDSKMPQVSLHFGTTHICGGL 325
QY 351 IDAQWVLTAAHCFVTRREKLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDI 410
DB 326 IDAQWVLTAAHCFVTRREKLEGMKYVAGTSLNHLQLEPAASIAEIIINSYTBEDDYDI 385
QY 411 ALMLRSKPLTSLAHIPACLPMHGQTFSLNETCMTITGFKTRTDDTSPFLREVQVNL 470
DB 386 ALMLRSKPLTSLAHIPACLPMHGQTFSLNETCMTITGFKTRTDDTSPFLREVQVNL 445
QY 471 DFKKCNVLYVDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGC 530
DB 446 DFKKCNVLYVDSYLTTPRMWCAQDLRGGRDSCQDSGGPLVCEQNNRWYLAGVTSWGTGC 505
QY 531 GQRNKPQVYTKVTEVLPWYISKMESEVFRKS 562
DB 506 GQRNKPQVYTKVTEVLPWYISKMESEVFRKS 537
```

RESULT 15
US-10-428-275-128
; Sequence 128, Application US/10428275
; Publication No. US20040067505A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-585
; CURRENT APPLICATION NUMBER: US/10/428,275
; PRIOR FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 09/966545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128514
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/569269
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/134315
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/619252
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185548
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 450
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 128
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-275-128

Query Match 92.5%; Score 2775.5; DB 4; Length 537;
Best Local Similarity 92.0%; Pred. No. 1.8e-164;
Matches 526; Conservative 0; Mismatches 1; Indels 45; Gaps 2;

QY 1 MERDSHGASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
Db 1 MERDSHGASPARTSPAGASPAQASPAQASPAQASPAQASPAQASPAQASPAQASPAQAS 60
QY 61 PAGTPGGRASPR-----ASPAQASPARASPALASLSRSSSGRSSSARASAVTTS 110
Db 61 PAGTPGGRASPRASPAQASPAQASPAQASPARASPALASLSRSSSGRSSSARASAVTTS 120
QY 111 PTRVYLVPATPVGAIVPIRSSPARSA PATRATRESPTSLPKFTWREGQQLPLICVLL 170
Db 121 PTRVYLVPATPVGAIVPIRSSPARSA PATRATRESPTSLPKFTWREGQQLPLICVLL 155
QY 171 IALVSLILFQWOGHTGIRYKEQRESCPKHVRCDGVNCKLSDELGCVRPMDKSL 230
Db 156 -----VQFMQHTGIRYKEQRESCPKHVRCDGVNCKLSDELGCVRPMDKSL 205
QY 231 LKIYSGSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFSLRYNSTI 290
Db 206 LKIYSGSSHQWLPICSSNMWDSYSEKTCQQLGFESAHRTTEVAHRDPFANSFSLRYNSTI 265
QY 291 QESLHRSCEPSORYISLQCSHGCLRAMTGRIVGALASDSKPMQVSLHFGTHICGTL 350
Db 266 QESLHRSCEPSORYISLQCSHGCLRAMTGRIVGALASDSKPMQVSLHFGTHICGTL 325
QY 351 IDAQWVLTAAHCFVTRKVLTEGMVYAGTSMHLQPEAASIAETIINSNYTDEEDDYDI 410
Db 326 IDAQWVLTAAHCFVTRKVLTEGMVYAGTSMHLQPEAASIAETIINSNYTDEEDDYDI 385
QY 411 ALMRISKPLTSAHHPACTLPMHGOTFSLNETCMTTGEGKRETDKTSPLREVOVNL 470
Db 386 ALMRISKPLTSAHHPACTLPMHGOTFSLNETCMTTGEGKRETDKTSPLREVOVNL 445
QY 471 DFKKCNVDYLVYDPSYLTPRMTCAGDLRGGRDSCGDSGGPLVCEQNNRWYLAGVTSWGTGC 530
Db 446 DFKKCNVDYLVYDPSYLTPRMTCAGDLRGGRDSCGDSGGPLVCEQNNRWYLAGVTSWGTGC 505

QY 531 GORNKPGVYTKVTEVLPWISKMESEVFRKS 562
Db 506 GORNKPGVYTKVTEVLPWISKMESEVFRKS 537

Search completed: September 16, 2006, 03:29:18
Job time : 182 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2006, 03:26:26 / Search time 35 Seconds

(without alignments)
1134.395 Million cell updates/sec

Title: US-10-806-370-12

Perfect score: 2999
Sequence: 1 MERSHGNASPARTPSAGAS.....TEVLPWYKMESEVRPRKS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 256596 seqs, 70647373 residues

Total number of hits satisfying chosen parameters: 256596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US09_NEW_PUB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US06_NEW_PUB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US07_NEW_PUB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US08_NEW_PUB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/1/pubppaa/PC1_NEW_PUB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US10_NEW_PUB.pep:*
- 7: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US11_NEW_PUB.pep:*
- 8: /EMC_Celerra_SIDS3/prodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729.5	24.3	453	US-10-196-749-64	Sequence 64, Appl
2	726	24.2	454	US-11-400-825-2	Sequence 2, Appl
3	694	23.1	432	US-10-196-749-310	Sequence 330, App
4	694	23.1	432	US-11-101-316-112	Sequence 112, App
5	608.5	20.3	416	US-11-376-673-112	Sequence 112, App
6	608.5	20.3	416	US-11-208-257-8	Sequence 8, Appl
7	543.5	18.1	305	US-11-293-697-3002	Sequence 3002, Ap
8	525.5	17.5	255	US-11-254-185-6	Sequence 6, Appl
9	525.5	17.5	255	US-11-254-185-38	Sequence 38, Appl
10	525.5	17.5	255	US-11-253-869-6	Sequence 6, Appl
11	525.5	17.5	255	US-11-253-869-38	Sequence 38, Appl
12	520.5	17.4	423	US-10-196-749-320	Sequence 320, App
13	520.5	17.4	423	US-11-101-316-106	Sequence 106, App
14	520.5	17.4	423	US-11-376-673-106	Sequence 106, App
15	518.5	17.3	853	US-10-530-187-242	Sequence 242, App
16	518.5	17.3	855	US-11-254-185-2	Sequence 2, Appl
17	518.5	17.3	855	US-11-253-869-2	Sequence 2, Appl
18	518.5	17.3	855	US-11-303-608-1	Sequence 1, Appl
19	515.5	17.2	311	US-11-327-490-41	Sequence 41, Appl
20	514	17.1	812	US-11-318-939-7	Sequence 7, Appl
21	510.5	17.0	343	US-11-359-554-3	Sequence 3, Appl
22	506	16.9	338	US-11-318-939-10	Sequence 10, Appl
23	506	16.9	331	US-10-522-668-2	Sequence 2, Appl
24	505	16.8	276	US-11-327-490-35	Sequence 35, Appl
25	498.5	16.6	339	US-11-404-745-2	Sequence 2, Appl

26	495	16.5	812	US-11-318-939-12	Sequence 12, Appl
27	493.5	16.5	790	US-11-350-703-1	Sequence 1, Appl
28	493.5	16.5	791	US-11-318-939-6	Sequence 6, Appl
29	493	16.4	333	US-11-318-939-8	Sequence 8, Appl
30	492.5	16.4	810	US-11-431-526-2	Sequence 2, Appl
31	492.5	16.4	810	US-11-431-663-2	Sequence 2, Appl
32	485.5	16.2	311	US-11-359-858-2	Sequence 2, Appl
33	480.5	16.0	262	US-11-359-554-2	Sequence 2, Appl
34	480	16.0	790	US-11-318-939-13	Sequence 13, Appl
35	479	16.0	245	US-11-254-185-36	Sequence 36, Appl
36	479	16.0	245	US-11-253-869-36	Sequence 36, Appl
37	479	16.0	275	US-10-530-798-24	Sequence 24, Appl
38	478	15.9	275	US-10-530-798-25	Sequence 25, Appl
39	476.5	15.9	272	US-11-327-490-37	Sequence 37, Appl
40	475.5	15.9	272	US-11-359-554-6	Sequence 6, Appl
41	473	15.8	275	US-10-530-798-26	Sequence 26, Appl
42	472.5	15.8	810	US-11-318-939-11	Sequence 11, Appl
43	470.5	15.7	272	US-11-327-490-40	Sequence 40, Appl
44	466	15.5	655	US-11-242-617-1	Sequence 1, Appl
45	465	15.5	275	US-10-530-798-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-10-196-749-64
Sequence 64, Application US/10196749
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ZHANG, ZEMIN
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT FILING DATE: 2002-07-16
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 64
LENGTH: 453
TYPE: PRT
ORGANISM: Homo Sapien
US-10-196-749-64

Query Match 24.3%; Score 729.5; DB 6; Length 453;
Best Local Similarity 37.8%; Pred. No. 6,8e-39;
Matches 158; Conservative 69; Mismatches 162; Indels 29; Gaps 10;

```
QY 161 LPL-----ICGVLILALVSLIILFQFMQHTGIRYKRESCPKHAYCQGVNDK 213
DB 44 LPLKEPPIIVIGIILITLALIGLGHFDC---SGKXRCRSSFICIELIARCDGVSDK 99
QY 214 LKSDDELGVRFMDLSLKIYSGSSHQWLPICSSNMNDSYSEKTCQOQGFESAHRTTEVA 273
DB 100 DQEDRYRCRVGGQNAVLQVFTLAS--WTKMCSDDMKHYANVACAQGLFESYSSDNIR 157
QY 274 -----HRDFANSFSLRYN--STIOBSLH--RSECPQRYISLQCSHCG--LRAMTGR1 321
DB 158 VSLRSGQFEFEFVSLIDHLLPDDKVTALHSHSVYRBCASGHVYTLQCTACGRRGYSRI 217
QY 322 VGGALASDSKMPWQVSLHGTTHICGGLIDAQWVLTAAHCFVTRREKVLBSMKVYAGTS 381
DB 218 VGGNNSLSQWPMQASLQFGYHLCCGSVITPLMTITTAHCYVDL--YLPSMTIQVGLV 275
QY 382 NLHOLPEAASIAE--IINSNTDEBDYDIALMRLSKPLTLSAHIHPACTLPMHGQFSLN 440
DB 276 SLDDNPASHLVKRYVYHSHKYPKRLGNDIALMKLAGPLTFEMIQPVCLPNSSENFPG 335
QY 441 ETCWITGFGKTRBTDDKTSPLREYQVNLIDFKCNDVLYVDSYLTTPRMWCAGDLRGGRD 500
DB 336 KVCMTSGMGATRDGD--ASPLVNHAAVPLISNKICNHRDVGGIISPSMLCAGYLTGGVD 394
QY 501 SCQSGSGPLVCEQNNRWTLAGVTSWGTCCGQRNRPYVTKTEVLPWYISKMESEVR 558
DB 395 SCQSGSGPLVCEQERRLMVLGATSFGLGCAEVNRPVYTRVTSFLDWIHQEMERDLK 452
```

RESULT 2

US-11-400-825-2
Sequence 2, Application US/11400825
Publication No. US20060177866A1
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
TITLE OF INVENTION: In Ovarian Carcinoma and Uses Thereof
FILE REFERENCE: D6192CIPD3
CURRENT APPLICATION NUMBER: US/11/400,825
CURRENT FILING DATE: 2006-04-07
PRIOR APPLICATION NUMBER: 10/455,720
PRIOR FILING DATE: 2003-06-05
NUMBER OF SEQ ID NOS: 154
SEQ ID NO 2
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: complete amino acid sequence of TADG-12
US-11-400-825-2

Query Match 24.2%; Score 726; DB 7; Length 454;
Best Local Similarity 37.7%; Pred. No. 1.1e-38;
Matches 154; Conservative 70; Mismatches 162; Indels 22; Gaps 9;

```
QY 164 ICGVLILALVSLIILFQFMQHTGIRYKRESCPKHAYCQGVNDCKLSDELACVR 223
DB 55 LGIILALIALAIGLGHFDC---SGKXRCRSSFICIELITRCDEVSCQKQEDRYRCVR 110
QY 224 FDMDKSLIKIYSGSSHQWLPICSSNMNDSYSEKTCQOQGFESAHRTTEVA-----HR 275
DB 111 VGGQNAVLQVFTLAS--WTKMCSDDMKHYANVACAQGLFESYSSDNIRVSLRSGQFRE 168
QY 276 DFANSFSLRYN--STIOBSLH--RSECPQRYISLQCSHCG--LRAMTGRIVGGALASDK 331
DB 169 EFVSLIDHLLPDDKVTALHSHSVYRBCASGHVYTLQCTACGRRGYSRIYGGNMSLSLQ 228
```

```
QY 332 WPMQVSLHFGTTHICGGLIDAQWVLTAAHCFVTRREKVLBSMKVYAGTSNLHOLPEAAS 391
DB 229 WPMQASLQFGHGLCGSSVITPLMTITTAHCYVDL--YLPSMTIQVGLVSLDDNPASH 286
QY 392 IAE-IINSNTDEBDYDIALMRLSKPLTLSAHIHPACTLPMHGQFSLNETCWTGTRGK 450
DB 287 LVEKIVYHSHKYPKRLGNDIALMKLAGPLTFEMIQPVCLPNSSENFPGDKVCMWISGWA 346
QY 451 TRETDKTSPLREYQVNLIDFKCNDVLYVDSYLTTPRMWCAGDLRGGRSCQSGSGPL 510
DB 347 TEDGD--ASPLVNHAAVPLISNKICNHRDVGGIISPSMLCAGYLTGGVDSQSGSGPL 405
QY 511 VCEQNNRWTLAGVTSWGTCCGQRNRPYVTKTEVLPWYISKMESEVR 558
DB 406 VQGERRLMKVLGATSFGLGCAEVNRPVYTRVTSFLDWIHQEMERDLK 453
```

RESULT 3

US-10-196-749-330
Sequence 330, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Auecin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 330
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapien

US-10-196-749-330

Query Match 23.1%; Score 694; DB 6; Length 432;
Best Local Similarity 34.9%; Pred. No. 1.1e-36;
Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;

```
QY 149 LPKRTWAGGQQLPLIGCVLILALVSLIILFQFMQHTGIRYKRESCPKH 202
DB 149 LPLKEPPIIVIGIILITLALIGLGHFDC---SGKXRCRSSFICIELIARCDGVSDK 99
```

```

Db 24 IPMETFR--KVGIPIIALLSLASIIIVVLIKVLKXYPLCGQ-----PLH 69
Qy 203 AV-----RCDGVNDCKLSDELGC-----VRPDMDSLKIKYSGSSHOMLPICSS 247
Db 70 FIPRKQCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129
Qy 248 NMNDYSEKTCQOLGFEASAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSRCPQRYISL 307
Db 130 NFEALAEFCRQNGYS---RAVEIGPDQDLDAVEITENSGELMRNNSGCLSGSLVSL 186
Qy 308 QCSHCGLRAMTGRVVGALASDSKMPWVSLHFGTTHICGTLIDAQVTLTAHCFEYTR 367
Db 187 HCLACGSLKTPRVVGGEEASVDSMPWVSIQYKQHVCGGSIIDPHVTLTAHCF--RK 244
Qy 368 EKVLEGMKVYAGTSLNHLQPEASIAEII---NSNTDEEDVDIALMRSLKPLTSAH 424
Db 245 HTDVFNKVRARGSDKLSFSLA--VAKITIIIEFNPWY---PKNDIDALMKQLPFTFSGT 300
Qy 425 IHPACLPNHGQTSINETCWITGFGKTRBDKTSPLFREYQVNLIDPKKNDVLYDSY 484
Db 301 VRPCLPFPEBELTPATPLWITGWFQKNGKSDILQASVQYIDSTRCNADDAVGE 360
Qy 485 LTPRMGAGDLRGGRDSCQDSDGGLVCEQNNRVYLAGVTSWGTCGCGRNKPGVYTKYTE 544
Db 361 VTERMMGAGIPBEGVDTCQDSDGGLM--YQSDQHVHVGIVSGVGCGLGPGSTPGVYTKVSA 419
Qy 545 VLPWYISKMESEV 557
Db 420 YLNMWYVWKAEL 432

```

RESULT 4

```

US-11-101-316-112
; Sequence 112, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101.316
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 112
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-112

```

Query Match 23.1%; Score 694; DB 7; Length 432;

Best Local Similarity 34.9%; Pred. No. 1,1e-36;

Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;

Qy 149 LPKFTWREGQKPLIGCVLLIALVLSLIL-----FQFWQHTGIRYKREGRSCPKH 202

Db 24 IPMETFR--KVGIPIIALLSLASIIIVVLIKVLKXYPLCGQ-----PLH 69

```

Qy 203 AV-----RCDGVNDCKLSDELGC-----VRPDMDSLKIKYSGSSHOMLPICSS 247
Db 70 FIPRKQCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129
Qy 248 NMNDYSEKTCQOLGFEASAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSRCPQRYISL 307
Db 130 NFEALAEFCRQNGYS---RAVEIGPDQDLDAVEITENSGELMRNNSGCLSGSLVSL 186
Qy 308 QCSHCGLRAMTGRVVGALASDSKMPWVSLHFGTTHICGTLIDAQVTLTAHCFEYTR 367
Db 187 HCLACGSLKTPRVVGGEEASVDSMPWVSIQYKQHVCGGSIIDPHVTLTAHCF--RK 244
Qy 368 EKVLEGMKVYAGTSLNHLQPEASIAEII---NSNTDEEDVDIALMRSLKPLTSAH 424
Db 245 HTDVFNKVRARGSDKLSFSLA--VAKITIIIEFNPWY---PKNDIDALMKQLPFTFSGT 300
Qy 425 IHPACLPNHGQTSINETCWITGFGKTRBDKTSPLFREYQVNLIDPKKNDVLYDSY 484
Db 301 VRPCLPFPEBELTPATPLWITGWFQKNGKSDILQASVQYIDSTRCNADDAVGE 360
Qy 485 LTPRMGAGDLRGGRDSCQDSDGGLVCEQNNRVYLAGVTSWGTCGCGRNKPGVYTKYTE 544
Db 361 VTERMMGAGIPBEGVDTCQDSDGGLM--YQSDQHVHVGIVSGVGCGLGPGSTPGVYTKVSA 419
Qy 545 VLPWYISKMESEV 557
Db 420 YLNMWYVWKAEL 432

```

RESULT 5

```

US-11-376-673-112
; Sequence 112, Application US/11376673
; Publication No. US20060160186A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
; TITLE OF INVENTION: LUNG TUMOR
; FILE REFERENCE: P3230R1C165C
; CURRENT APPLICATION NUMBER: US/11/376.673
; PRIOR FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: 10/063742
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 112
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-376-673-112

```

Query Match 23.1%; Score 694; DB 7; Length 432;

Best Local Similarity 34.9%; Pred. No. 1,1e-36;

Matches 151; Conservative 73; Mismatches 161; Indels 48; Gaps 11;

Qy 149 LPKFTWREGQKPLIGCVLLIALVLSLIL-----FQFWQHTGIRYKREGRSCPKH 202

Db 24 IPMETFR--KVGIPIIALLSLASIIIVVLIKVLKXYPLCGQ-----PLH 69

Qy 203 AV-----RCDGVNDCKLSDELGC-----VRPDMDSLKIKYSGSSHOMLPICSS 247

Db 70 FIPRKQCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTLQVLDATGNMFSACFD 129

Qy 248 NMNDYSEKTCQOLGFEASAHRTTEVAHRDPFANSFSLRYNSTIOESLHRSRCPQRYISL 307

Db 130 NFEALAEFCRQNGYS---RAVEIGPDQDLDAVEITENSGELMRNNSGCLSGSLVSL 186

Qy 308 QCSHCGLRAMTGRVVGALASDSKMPWVSLHFGTTHICGTLIDAQVTLTAHCFEYTR 367

Db 187 HCLACGSLKTPRVVGGEEASVDSMPWVSIQYKQHVCGGSIIDPHVTLTAHCF--RK 244

Qy 368 EKVLEGMKVYAGTSLNHLQPEASIAEII---NSNTDEEDVDIALMRSLKPLTSAH 424

Db 245 HTDVFNKVRARGSDKLSFSLA--VAKITIIIEFNPWY---PKNDIDALMKQLPFTFSGT 300

Qy 425 IHPACLPNHGQTSINETCWITGFGKTRBDKTSPLFREYQVNLIDPKKNDVLYDSY 484

Db 301 VRPCLPFPEBELTPATPLWITGWFQKNGKSDILQASVQYIDSTRCNADDAVGE 360

Qy 485 LTPRMGAGDLRGGRDSCQDSDGGLVCEQNNRVYLAGVTSWGTCGCGRNKPGVYTKYTE 544

Db 361 VTERMMGAGIPBEGVDTCQDSDGGLM--YQSDQHVHVGIVSGVGCGLGPGSTPGVYTKVSA 419

Qy 545 VLPWYISKMESEV 557

Db 420 YLNMWYVWKAEL 432

```

Db      130 NFEALAEATACRQWQYS---RAVEIGPDODLVEITENSQELRMNRSSGCLSGSLVSL 186
Qy      308 QCSHCGELRAMTGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAAHCFEYTR 367
Db      187 HCLACGKSLKTPRVVGGSEBASVDSPQVSIQYDQHQVCGSSILDPHNVLTAAHCF--RK 244
Qy      368 EKVLGKRVVAGTSMHLQLEPAASIAEIIII--NSNYTDEEDDYDIALMRLSKPLTSLAH 424
Db      245 HTDVFNMKVRAQSDKLGSPSLA-VAKIIIIIEFNMY--PRONDIALMKLQPLTFSGT 300
Qy      425 IHPACLPMHGQTFSLNFTCMINGPQKTRBDKTSPLREVOYNLIDPKKCNVDLYVDSY 484
Db      301 VRPLCPFEDELTATPLMTIIGMGTQKONGKMSDILQASVQVLDSTRCNADDAVQGE 360
Qy      485 LTPRMACAGDLRGGRDSGCGSGPLVCEQNNRWYLAQVTSWGTGCGQRNKPVTYKYTE 544
Db      361 VTEKMKMGAGIBGGVDTQCGDSGGLM-YQSPQMHVIGVSVNGYCGGSPSTGYTTKXSA 419
Qy      545 VLPWYTSKMESEV 557
Db      420 YLNMWYVWKAEL 432

```

```

RESULT 6
US-11-208-257-8
; Sequence 8, Application US/11208257
; Publication No. US20060101531A1
; GENERAL INFORMATION:
; APPLICANT: Vasiloukhin et al., Valeri
; TITLE OF INVENTION: ANIMAL MODELS OF CANCER DEVELOPMENT AND METASTASIS
; FILE REFERENCE: 14538A-008510US
; CURRENT APPLICATION NUMBER: US/11/208,257
; CURRENT FILING DATE: 2005-08-19
; PRIOR APPLICATION NUMBER: US 60/603,043
; PRIOR FILING DATE: 2004-08-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 8
; LENGTH: 416
; TYPE: PRT
; ORGANISM: mouse
US-11-208-257-8

```

```

Query Match      20.3%; Score 608.5; DB 7; Length 416;
Best Local Similarity 36.8%; Pred. No. 2.7e-31;
Matches 140; Conservative 60; Mismatches 135; Indels 45; Gaps 12;

Qy      209 VVDCLEKDE--LGCYRFDMDKSLIKIYSGSSHQMLPICSSNMNDYSSEKTCQQLGFESA 266
Db      38  IVTLLQSDQBLPYQVQLSPGDSRLAVLDKTEGTRLLCSSRSNARVAGLGCSEMGFLRA 97
Qy      267 HRTVEARDF-----ANSFS-----ILRYNTIOESLHRECPRSORYISLQCSHC 312
Db      98  -----LHSELDVRYTAGANGTSGFCVDEBGLPLAQLRLIDVSYVDCRCRGRFLITTCDDC 152
Qy      313 GLRAM-TGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLTAAHCFEYTRKY 371
Db      153 GRRKLPRVRIYQCGQSSLSGRWPQVSLAYDTHLCCGSLSDQVLTAAHC-FPERNRVL 211
Qy      372 EGMKYYAG---TSLHLQLEPA--SIAEIIINSY-----TDEEDDYDIALMRLSPL 419
Db      212 SRWRVFAQAVARTS-----PRAVOLGVQAVIYHGGLPFRDPTDIDENSNDIALVHLSL 266
Qy      420 TLSAIIHACLPMHGQTFSLNFTCMITGFKTRBDKTSPLREVOYNLIDFKKCNVDL 479
Db      267 PLTEIYQVCLPAQAGALVDGKVCITVQWNT-QYTGQOAMVLOARPIISNEVCNSPD 325
Qy      480 VYDSYLRPMACAGDLRGGRDSGCGSGPLVCEQN---NRWYLAQVTSWGTGCGQRN 535
Db      326 FYGNQIKPKMKFCAGYPEGGIDACQDSGSPFVCEDSISGTSRMRLCGIVSMOTGALARK 385
Qy      536 PGVYTKTEVLPWYTSKMESEV 555

```

```

Db      386 PGVYTKVTFDEWIFKAIRT 405

```

```

RESULT 7
US-11-293-697-3002
; Sequence 3002, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; NUMBER OF SEQ ID NOS: 2002-03-28
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3002
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3002

```

```

Query Match      18.1%; Score 543.5; DB 7; Length 305;
Best Local Similarity 40.9%; Pred. No. 2.5e-27;
Matches 113; Conservative 40; Mismatches 96; Indels 27; Gaps 8;

Qy      299 CPSQRYSILQCSHCGELRAM-TGRIVGALASDSKMPQVSLHFGTTHICGGTLIDAQWVLT 357
Db      33  CPTD-----GRRRLPYDRIVGGKRDTSLGRMPQVSLRYDGAHLCCGSLSGDVL 83
Qy      358 TAAHCFEYTRKRVLEGWVYAGT--SNLHQLPEAASIAEIIINSY-----TDEEDDY 408
Db      84  TAAHC-FPERNRVLSRWRVFAQAVAAQASPHGL--QLGVQAVVYHGCVLPFRDPNSEKSN 140
Qy      409 DIALMRLSKPLTSLAHHPACLPMHGQTFSLNFTCMITGFKTRBDKTSPLREVOYN 468
Db      141 DIALVHLSPLPLTEYIQPVCLPAAGALVDGKICTVGMONTQYVGGQAG-VLOEARVP 199
Qy      469 LIDFKKCNVDLYVQSYLTPRMACAGDLRGGRDSGCGSGPLVCE---QNNRWYLAQVT 524
Db      200 IISNDVCAGADPTGNQIRPKMFCAGYPEGGIDACQDSGSPFVCEDSISRTPRRLCGIV 259
Qy      525 SWGTGCGQRNKPQVYTKTEVLPWYTSKMESEVFR 560
Db      260 SWGTGALAQKPGVYTKVSDEREMIFQAIKXVGR 295

```

```

RESULT 8
US-11-254-185-6
; Sequence 6, Application US/11254185
; Publication No. US20060099625A1
; GENERAL INFORMATION:
; APPLICANT: CRAIK, CHARLES S.
; APPLICANT: TAKEUCHI, TOSHIHIKO
; TITLE OF INVENTION: MT-SPL SERINE PROTEASE
; FILE REFERENCE: 28644-701.302
; CURRENT APPLICATION NUMBER: US/11/254,185
; CURRENT FILING DATE: 2005-10-18
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 6
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Protein fragment/domain
US-11-254-185-6

```

```

Query Match      17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

```

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLSGRMPQVSLRYDGAHLCCGSLSGDWVLTAAHC-FPERNRVLSRMRVFAGA 59
QY 381 ---SNLHQLPEAASIAEIIINSNY-----TDEEDVDYIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIOPVCLP 117
QY 432 MHGQFSLNETCMTWTFGKTRTDKTSPLREVOVNLIDFKKNDVLYVDSYLTTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADFYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTCCGGRNRPVYTKVTEVLP 547
DB 177 AGYPEGGIDACQSGSGPLVECDISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDERE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 9
US-11-254-185-38
Sequence 38, Application US/11254185
Publication No. US20060099625A1
GENERAL INFORMATION:
APPLICANT: CRAIK, CHARLES S.
APPLICANT: TAKEUCHI, TOSHIHIKO
TITLE OF INVENTION: MT-SPI SERINE PROTEASE
FILE REFERENCE: 28644-701.302
CURRENT APPLICATION NUMBER: US/11/254.185
CURRENT FILING DATE: 2005-10-18
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn version 3.3
SEQ ID NO 38
LENGTH: 255
TYPE: PRT
ORGANISM: Homo sapiens
US-11-254-185-38

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLSGRMPQVSLRYDGAHLCCGSLSGDWVLTAAHC-FPERNRVLSRMRVFAGA 59
QY 381 ---SNLHQLPEAASIAEIIINSNY-----TDEEDVDYIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIOPVCLP 117
QY 432 MHGQFSLNETCMTWTFGKTRTDKTSPLREVOVNLIDFKKNDVLYVDSYLTTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADFYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTCCGGRNRPVYTKVTEVLP 547
DB 177 AGYPEGGIDACQSGSGPLVECDISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDERE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 10
US-11-253-869-6
Sequence 6, Application US/11253869
Publication No. US20060104979A1
GENERAL INFORMATION:
APPLICANT: CRAIK, CHARLES S.

APPLICANT: TAKEUCHI, TOSHIHIKO
APPLICANT: SCHUMAN, MARC
TITLE OF INVENTION: MT-SPI POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: 28644-701.303
CURRENT APPLICATION NUMBER: US/11/253.869
CURRENT FILING DATE: 2005-10-18
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn version 3.3
SEQ ID NO 6
LENGTH: 255
TYPE: PRT
ORGANISM: Artificial
FEATURES:
OTHER INFORMATION: Protein fragment/domain
US-11-253-869-6

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLSGRMPQVSLRYDGAHLCCGSLSGDWVLTAAHC-FPERNRVLSRMRVFAGA 59
QY 381 ---SNLHQLPEAASIAEIIINSNY-----TDEEDVDYIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIOPVCLP 117
QY 432 MHGQFSLNETCMTWTFGKTRTDKTSPLREVOVNLIDFKKNDVLYVDSYLTTPRMC 491
DB 118 AAGQALVDGKICTYTGWNTQYGGQAG-VLQEARVPIISNDVCGADFYGNQIKPKMFC 176
QY 492 AGDLRGGRDSCQSGSGPLVE---QNNRWYLAGVTSWGTCCGGRNRPVYTKVTEVLP 547
DB 177 AGYPEGGIDACQSGSGPLVECDISRTPRWRLCGIVSWGTCALAQKRGVYTKVSDERE 236
QY 548 WYSKMES 555
DB 237 WIFQAIKT 244

RESULT 11
US-11-253-869-38
Sequence 38, Application US/11253869
Publication No. US20060104979A1
GENERAL INFORMATION:
APPLICANT: CRAIK, CHARLES S.
APPLICANT: TAKEUCHI, TOSHIHIKO
TITLE OF INVENTION: MT-SPI POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: 28644-701.303
CURRENT APPLICATION NUMBER: US/11/253.869
CURRENT FILING DATE: 2005-10-18
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn version 3.3
SEQ ID NO 38
LENGTH: 255
TYPE: PRT
ORGANISM: Homo sapiens
US-11-253-869-38

Query Match 17.5%; Score 525.5; DB 7; Length 255;
Best Local Similarity 42.3%; Pred. No. 2.9e-26;
Matches 105; Conservative 38; Mismatches 88; Indels 17; Gaps 6;

QY 321 IVGGALASDSKMPQVSLHFGTTHICGGLTIDAQWVLTAAHCFVTRKVLGKVVYAGT 380
DB 1 IVGGRDTSLSGRMPQVSLRYDGAHLCCGSLSGDWVLTAAHC-FPERNRVLSRMRVFAGA 59
QY 381 ---SNLHQLPEAASIAEIIINSNY-----TDEEDVDYIALMRSLKPLTSLAHHPACLP 431
DB 60 VAQASPHGL--QLGVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLPLEYIOPVCLP 117

Qy 432 MHGQFSINETCMTGFGKTRTBDKTSPLREVOVNLIDFKKNDYLVYDSYLTFRMNC 491
Db 118 AAGQALVQDKICTVGMWNTQYGGQAG-VLQEARVPIISNDVCGADFGYQIKPKHFC 176
Qy 492 AGDLAGSDSCGDSGGPLVCE---QNNRWYLAGVTSWGTCCGQNRPGVYTKTEVLP 547
Db 177 AGYPRGGIDACGSDGSPVCEBDSISRTPRMVLGIVSGTGCALAKPKGYVTKVSDPRE 236
Qy 548 WYSPKMS 555
Db 237 WIFQAIKT 244

RESULT 12
US-10-196-749-320

/ Sequence 320, Application US/10196749
/ Publication No. US20060094864A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3430R1C340
/ CURRENT FILING DATE: 2002-07-16
/ PRIOR FILING DATE: 2002-07-16
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 320
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-196-749-320

Query Match 17.4%; Score 520.5; DB 6; Length 423;
Best Local Similarity 39.6%; Pred. No. 1e-25;
Matches 109; Conservative 41; Mismatches 112; Indels 13; Gaps 5;

Qy 279 NSFSLRNSTIOSSLRHSECPQRYISLQSHCGLRMTGRIYCGALASDKMPQVSL 338
Db 158 HSKIKIKINKETDYSYLNHCCGTRSKTL-----GQSLRIYGVTEVEGEMWQASL 209
Qy 339 HFGTHICGTLIDQAWYLTAAHCFVYRREKVLBGWKYVAGTSNHLQPEASIAIEIIN 398
Db 210 QMDGSHRCGATLINATWVSAHCF--TTYKNPARWTASFGVT-IKPSKMRGRLRIIVH 266

Qy 399 SNYTDDEDDYDIALMRLSKPLTSLAHIPACLPMHGQFSINETCMTGFGKTRTBDKT 458
Db 267 EKYKHPSHDYDISLAELSSPVYTNNAVHRCVLPDASVEFGQDWNFVYGFALK-NDGYS 325
Qy 459 SPFLREVOVNLIDFKKNDYLVYDSYLTFRMNCAGDLRGSDSCGDSGGPLV-CEQNNR 517
Db 326 QNHLRQAGVTLIDATTCTNCPQAYNDALTPRMLCAGSLGKTDACGSDGGLVSSDARDI 385
Qy 518 WYLAGVTSWGTCCGQNRKPGVYTKTEVLPWYISK 552
Db 386 WYLAGVTSWGDCAKPKRPGVYTRVYALRDWITSK 420

RESULT 13
US-11-101-316-106

/ Sequence 106, Application US/11101316
/ Publication No. US20060099657A1
/ GENERAL INFORMATION:
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
/ TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
/ FILE REFERENCE: P3230R1C1C1
/ CURRENT APPLICATION NUMBER: US/11/101,316
/ PRIOR FILING DATE: 2005-04-06
/ PRIOR APPLICATION NUMBER: 10/063526
/ PRIOR FILING DATE: 2002-05-03
/ PRIOR APPLICATION NUMBER: 10/006867
/ PRIOR FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: PCT/US00/23328
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: 09/380137
/ PRIOR FILING DATE: 1999-08-25
/ PRIOR APPLICATION NUMBER: PCT/US99/12252
/ PRIOR FILING DATE: 1999-06-02
/ PRIOR APPLICATION NUMBER: 60/087759
/ PRIOR FILING DATE: 1998-06-02
/ NUMBER OF SEQ ID NOS: 170
/ SEQ ID NO 106
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-11-101-316-106

Query Match 17.4%; Score 520.5; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 1e-25;
Matches 109; Conservative 41; Mismatches 112; Indels 13; Gaps 5;

Qy 279 NSFSLRNSTIOSSLRHSECPQRYISLQSHCGLRMTGRIYCGALASDKMPQVSL 338
Db 158 HSKIKIKINKETDYSYLNHCCGTRSKTL-----GQSLRIYGVTEVEGEMWQASL 209
Qy 339 HFGTHICGTLIDQAWYLTAAHCFVYRREKVLBGWKYVAGTSNHLQPEASIAIEIIN 398
Db 210 QMDGSHRCGATLINATWVSAHCF--TTYKNPARWTASFGVT-IKPSKMRGRLRIIVH 266
Qy 399 SNYTDDEDDYDIALMRLSKPLTSLAHIPACLPMHGQFSINETCMTGFGKTRTBDKT 458
Db 267 EKYKHPSHDYDISLAELSSPVYTNNAVHRCVLPDASVEFGQDWNFVYGFALK-NDGYS 325
Qy 459 SPFLREVOVNLIDFKKNDYLVYDSYLTFRMNCAGDLRGSDSCGDSGGPLV-CEQNNR 517
Db 326 QNHLRQAGVTLIDATTCTNCPQAYNDALTPRMLCAGSLGKTDACGSDGGLVSSDARDI 385
Qy 518 WYLAGVTSWGTCCGQNRKPGVYTKTEVLPWYISK 552
Db 386 WYLAGVTSWGDCAKPKRPGVYTRVYALRDWITSK 420

RESULT 14
US-11-376-673-106
; Sequence 106, Application US/11376673
; Publication No. US20060160186A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
; FILE REFERENCE: P3230R1C165C
; CURRENT APPLICATION NUMBER: US/11/376,673
; PRIOR FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: 10/063742
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 106
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-376-673-106

Query Match 17.4%; Score 520.5; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 1e-25;
Matches 109; Conservative 41; Mismatches 112; Indels 13; Gaps 5;
QY 279 NSFSILRYNSTIQLSHRSECPQRYISLQCSHGLRAMTSRIVGALASDSKPMQVSL 338
Db 158 HSVKIKINKETSTSYLNHCCTGRSKTL-----QGSRLRYGQTEVEGEMPMQSL 209
QY 339 HFGTTHICGGTLIDAGVLTAAHCFEFTVREKYLEGKKYVAGTSNHLPEAASIAEIIIN 398
Db 210 QMDSHRCAGATLINAATVLSAHCFT--TTYKNPARMTASFGVT--IKSKMKRGLRRIIVH 266
QY 399 SNTYDEEDVDYIALMRSLPILSAHIHPACTPMHGQTFSLNETCMTITGFGKTRTDDKT 458
Db 267 EKYGHPSHDYISLAEISSPVPTNAVHRCVLPDASIEFGQDVMFTVGFALK-NDGYS 325
QY 459 SPFLREVQVNLIDPKCNDYLVDSYLTFRMTCAGDLRGHDSQGDGGGPLY-CEQNNR 517
Db 326 QNHRLQAQVTLIDATTCNEPQAYNDALTPRMLCAGSLGKTDACQGDGGGPLYSSDARDI 385
QY 518 WYLAGVTSWGTGCGQRNKGYYTYKTEVLPWYISK 552
Db 386 WYLAGIVSWGDECAKPNKPGYTYTVALRDWITSK 420

RESULT 15
US-10-530-187-242
; Sequence 242, Application US/10530187
; Publication No. US20060183120A1
; GENERAL INFORMATION:
; APPLICANT: THE Bin Tean
; APPLICANT: TAKAHASHI, Masayuki
; TITLE OF INVENTION: Molecular Subclassification of Kidney Tumors and the Discovery of
; FILE REFERENCE: 28927-0014
; CURRENT APPLICATION NUMBER: US/10/530,187
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US2003/031476
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: US 60/415,775
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 242
; LENGTH: 853
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-530-187-242

Query Match 17.3%; Score 518.5; DB 6; Length 853;
Best Local Similarity 32.5%; Pred. No. 2.8e-25;
Matches 135; Conservative 50; Mismatches 137; Indels 93; Gaps 17;

QY 199 CPKIAVDCGVNDCKLSDELICVRFPMDSLSLKIYSGSSHQ-----WLPICS 246
Db 462 CIRKELRCDDGADCTDHSDELNC-----SCDAGHQFTCKXKFCRPLFW--VCD 507
QY 247 S--NMNDSYSEKTC-----QQGPFESAHRTTEVAHRFANFSILR 285
Db 508 SYNDGNSDBEQGSCPAQTFRCNSGKLSKQQCNKNDGCGSDSDASCPKVVVYCTK 567
QY 286 YNSTIQESLHRS-----ECPQRYISLQCS-----HCGLRAMT--GRIVGALASDSK 331
Db 568 HTYRCLANGCLSKNPECDGKE---DCSDGSDKDCDCGLRSTTRQARVVGTDADGCE 623
QY 332 WPMQVSLH-FGTTHICGTLIDAGVLTAAHCFEFTVREKYLEG-----WKYVAGTSN 382
Db 624 WPMQVSLHALGQGHICGASLISPMVLVSAHCYIDR-----GFRYSDPTQWTAFLG--- 675
QY 383 LHQLPEAAS-----IAEIIINSNTYDEEDVDYIALMRSLPILSAHIHPACTPMHGQ 435
Db 676 LHDQSQRSAPQVQERRRLKRIISHPFNDFTFDYDIALLELEKPAEYSSMVARPCLPDASH 735
QY 436 TFSLETCTMTGFGKTRTDDKTSPFLREVQVNLIDPKCNDYLVDSYLTFRMTCAGDL 495
Db 736 VFPAGKAIWYTGWHT--QYGTGALILQKGEIRIYINOTTENL--PQGITPRMTCGFL 792
QY 496 RGRHDSQGDGGGPLY-VCEQNNRWYLAGVTSWGTGCGQRNKGYYTYKTEVLPWI 549
Db 793 SGAVDSQGDGGGPLYSSVEADGRIFQAGVSWGCGCAQRNKGYYTYTVALPFRDWI 847

Search completed: September 16, 2006, 03:29:58
Job time : 36 secs

This Page Blank (usp10)